ARTICLE IN PRESS

Available online at www.sciencedirect.com

ScienceDirect

Systematic and Applied Microbiology ▮ (▮▮▮▮) ▮▮▮–▮▮▮

SYSTEMATIC AND APPLIED MICROBIOLOGY

www.elsevier.de/syapm

Phylogeny and identification of *Pantoea* species associated with plants, humans and the natural environment based on multilocus sequence analysis (MLSA) ☆

Carrie Brady[a], Ilse Cleenwerck[b], Stephanus Venter[a], Marc Vancanneyt[b], Jean Swings[b], Teresa Coutinho[a,*]

[a]*Department of Microbiology and Plant Pathology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria 0002, South Africa*

[b]*BCCM/LMG Bacteria Collection, Ghent University, K.L. Ledeganckstraat 35, B-9000 Ghent, Belgium*

Received 26 May 2008

Abstract

Species belonging to the genus of *Pantoea* are commonly isolated from plants, humans and the natural environment. The species of the genus are phenotypically closely related, making rapid identification of *Pantoea* strains to the species level difficult. Multilocus sequence analysis (MLSA) was evaluated as a means for rapid classification and identification of *Pantoea* strains. Four housekeeping genes, *gyrB*, *rpoB*, *atpD* and *infB*, were sequenced for strains assigned to the genus. Included in the study were (1) reference strains from the seven currently recognized species of *Pantoea*, (2) strains belonging to Brenner DNA groups II, IV and V, previously isolated from clinical samples and difficult to identify because of high phenotypic similarity to *P. agglomerans* or *P. ananatis* and (3) isolates from diseased *Eucalyptus*, maize and onion, assigned to the genus on the basis of phenotypic tests. Phylogenetic trees were constructed from the sequences of the four housekeeping genes. The ''core'' *Pantoea* species formed a cluster separate from the ''Japanese'' species which formed a tight cluster that included the genus *Tatumella* when the tree was based on concatenated sequences of the four genes. The MLSA data further suggested the existence of ten potential novel species, phylogenetically related to the currently recognized *Pantoea* species and the possible inclusion of *Pectobacterium cypripedii* in the genus *Pantoea*. When compared with DNA–DNA hybridization data, a good congruence was observed between both methods, with *gyrB* sequence data being the most consistent. In conclusion, MLSA of partial nucleotide sequences of the genes *gyrB*, *rpoB*, *atpD* and *infB* can be used for classification, identification and phylogenetic analyses of *Pantoea* strains.

Keywords: *Pantoea*; MLSA; Phylogeny; Identification; Multigene

☆ *Note*: The GenBank/EMBL accession numbers for the sequences presented in this study are: EF988753–EF988838, EU145260–EU145275, EU344757–EU344760, FJ187830–FJ187834 (*gyrB* gene); EF988925–EF989010, EU145292–EU145307, EU344765–EU344768, FJ187840–FJ187844 (*rpoB* gene); EF988667–EF988752, EU145244–EU145259, EU344753–EU344756, FJ187825–FJ187829 (*atpD* gene); EF988839–EF988924, EU145276–EU145291, EU344761–EU344764, FJ187835–FJ187839 (*infB* gene) and EF688006–EF688012, EU216734–EU216737, EU344769–EU344770 (16S rRNA).

*Corresponding author. Tel.: +2712 420 3934; fax: +2712 420 3960.
E-mail address: teresa.coutinho@fabi.up.ac.za (T. Coutinho).

0723-2020/$ - see front matter © 2008 Elsevier GmbH. All rights reserved.
doi:10.1016/j.syapm.2008.09.004

Please cite this article as: C. Brady, et al., Phylogeny and identification of *Pantoea* species associated with plants, humans and the natural environment based on multilocus sequence analysis..., Syst. Appl. Microbiol. (2008), doi:10.1016/j.syapm.2008.09.004

Introduction

The genus *Pantoea* belongs within the family *Enterobacteriaceae* and was proposed by Gavini et al. [17] for two groups of strains that were, at that time, assigned to the *Erwinia herbicola–Enterobacter agglomerans* complex. This complex covered many phena [16,44] and genomic groups [3], some of which were later designated as new genera [19]. The genus *Pantoea* comprises at present seven validly published species, namely *Pantoea agglomerans* and *P. dispersa* [17], *P. citrea*, *P. punctata* and *P. terrea* [23] and *P. ananatis* and *P. stewartii* [30]. However, Grimont and Grimont [19] stated that the genus *Pantoea* can be envisioned to include DNA groups I, II, IV and V as determined by Brenner et al. [3]. It was further observed that the species *P. citrea*, *P. punctata* and *P. terrea*, isolated in Japan and described by Kageyama et al. [23] differed from the "core" *Pantoea* species in several biochemical or nutritional characteristics. Grimont and Grimont [19] determined the phylogenetic position of all currently recognized *Pantoea* species and DNA groups of Brenner et al. [3] using 16S rRNA- and *rpoB*-sequence comparisons and found that the "Japanese" species constituted a cluster that joined the *Pantoea* cluster at a lower level. They concluded that more taxonomic work was needed to justify the assignment of these species to the genus *Pantoea*.

Several species belonging to the genus *Pantoea* are known as plant pathogens. Stewart's vascular wilt is a disease of sweet corn and maize caused by *Pantoea stewartii* subsp. *stewartii* [40], *P. agglomerans* causes crown and root gall disease of gypsophila and beet [4,7] and *P. ananatis* causes a variety of diseases on a wide range of hosts including bacterial blight and dieback of *Eucalyptus* [9], stem necrosis of rice [8] and brown stalk rot of maize [18]. Recently, *Pantoea* strains were isolated from young *Eucalyptus* trees in Uganda, Argentina and Uruguay, showing a disease similar to bacterial blight and dieback of *Eucalyptus* in South Africa. *Pantoea* strains, not belonging to *P. ananatis* and causing brown stalk rot of maize were also isolated. Similarly, *Pantoea* strains that could not clearly be identified to the species level were isolated from onion in South Africa and the USA. All of these isolates were assigned to the genus *Pantoea* on the basis of phenotypic tests.

All strains of DNA groups I, II, IV and V determined by Brenner et al. [3] were isolated from human sources. In recent years *Pantoea* strains have been consistently linked with human infections [12,15,26,28,36,43]. The increasing isolations of *Pantoea* strains from the natural environment, from human infections and diseased plant material highlights the need for a technique that enables fast and reliable classification and identification of *Pantoea* strains worldwide.

Partial sequences of protein-encoding genes have been proven useful for species identification and as phylogenetic markers in the family *Enterobacteriaceae*. The following genes have been evaluated for these purposes: *rpoB* [13,24,27,31,39], *infB* [22], *groEL* [29]; *gyrB* [11], *tuf* and *atpD* [33,46], *carA* and *recA* [46]. Results have shown that they are more reliable than 16S rRNA gene sequences for species identification and for determining intra- and some inter-generic relationships [11,22,31,33]. For determination of phylogenetic relationships, it has been advised to use sequence data from more than one gene [11,25], to reduce the possibility of ambiguities caused by genetic recombination or specific selection [22]. 16S rRNA gene sequences, on the other hand, appeared to be useful for determination of phylogenies between distantly related *Enterobacteriaceae* [11].

In the present study, the phylogeny of all validly published species of *Pantoea*, DNA groups II, IV and V of Brenner et al. [3], which belong to the genus *Pantoea* according to Grimont and Grimont [19], as well as isolates assigned to *Pantoea* on the basis of phenotypic tests, were investigated using multilocus sequence analysis (MLSA) of the protein-encoding genes *gyrB*, *rpoB*, *atpD* and *infB*. These genes encode DNA gyrase, RNA polymerase β subunit, ATP synthase β subunit and initiation translation factor 2, respectively.

Material and methods

Strains investigated and DNA extraction

A detailed summary of the strains used in this study is listed in Table 1. All strains used in this study are maintained in the Bacterial Culture Collection (BCC) of the Forestry and Agricultural Biotechnology Institute (FABI) and representative isolates have been deposited in the BCCM/LMG Collection, University of Ghent, Belgium. Care was taken to include type strains and reference strains for which existing DNA–DNA hybridization data are available. The identities of strains included in this study were confirmed using FAFLP analysis as previously described [2].

Genomic DNA was extracted from each of the bacterial strains using an alkali extraction method [32] and stored at −20 °C.

Primer design

External primers for amplification and internal primers for sequencing were designed based on sequence alignments of strains representative of multiple species belonging to the family *Enterobacteriaceae*. The primers used in previous studies of the *Enterobacteriaceae* [11,22,31,33], were the basis for the design of the primers

Please cite this article as: C. Brady, et al., Phylogeny and identification of *Pantoea* species associated with plants, humans and the natural environment based on multilocus sequence analysis..., Syst. Appl. Microbiol. (2008), doi:10.1016/j.syapm.2008.09.004

Table 1. Strains used in this study

Species name	Strain no.	Source	Place of isolation	% DNA[a] binding to *P. agglomerans*	% DNA[b] binding to *P. ananatis*	% DNA[c] binding to *T. ptyseos*	% DNA[d] binding in MLSA groups
Pantoea agglomerans	LMG 1286^T	Human	Zimbabwe	100	21		
	LMG 2554	Scarlet runner bean	UK				
	LMG 2565	Cereal	Canada	93			
	LMG 2572	Wheat	Canada				
	LMG 2596	Onion	South Africa	93			
	LMG 2660	Wisteria	Japan	93			
	SUH 2 [syn. LMG 2596]	Onion	South Africa				
Pantoea agglomerans pv. *gypsophilae*	LMG 2553	Gypsophila	Unknown				
Pantoea agglomerans pv. *betae*	BCC 734	Beet	Unknown				
Pantoea ananatis	LMG 2665^T	Pineapple	Brazil	21	100		
	LMG 2668	Pineapple	Hawaii				
	LMG 2676	*Puccinia graminis*	USA				
	LMG 2678	*Puccinia graminis*	Zimbabwe				
	LMG 20103	*Eucalyptus*	South Africa				
	LMG 20104	*Eucalyptus*	South Africa				
	LMG 20106	*Eucalyptus*	South Africa				
	BCC 114	*Eucalyptus*	South Africa				
	BCC 150 = ATCC 35400	Honeydew melon	USA				
	LMG 24190 = R-27854	Onion	USA		90		
	LMG 24193 = R-27860	Onion seed	South Africa		92		
	BD 333	Onion seed	South Africa				
	BD 336	Onion seed	South Africa				
	LMG 24191 = R-27858	Maize	South Africa		90		
	LMG 24192 = R-27859	Maize	South Africa		93		
	BD 561	Maize	South Africa				
	BD 577	Maize	South Africa				
	BD 588	Maize	South Africa				
	BD 602	Maize	South Africa				
	BD 622	Maize	South Africa				
	BD 640	Maize	South Africa				
	BD 647	Maize	South Africa				
Pantoea stewartii ssp. *stewartii*	LMG 2715^T	Corn	USA	6	20		
	LMG 2713	Corn	USA				
	LMG 2718	Corn	USA				
Pantoea stewartii ssp. *indologenes*	LMG 2632^T	Fox millet	India				
	LMG 2630	Guar gum powder	Unknown				
	LMG 2631	Millet	India				
	LMG 2671	Pineapple	Hawaii				
	LMG 2673	Pineapple	Hawaii				
	BCC 099	Sudangrass	USA				
	BCC 118	*Eucalyptus*	South Africa				

Please cite this article as: C. Brady, et al., Phylogeny and identification of *Pantoea* species associated with plants, humans and the natural environment based on multilocus sequence analysis..., Syst. Appl. Microbiol. (2008), doi:10.1016/j.syapm.2008.09.004

Table 1. (*continued*)

Species name	Strain no.	Source	Place of isolation	% DNA[a] binding to *P. agglomerans*	% DNA[b] binding to *P. ananatis*	% DNA[c] binding to *T. ptyseos*	% DNA[d] binding in MLSA groups
Pantoea dispersa	LMG 2603^T	Soil	Japan	24	20		
	LMG 2602	Sorghum	India				
	LMG 2604	Wild rose	Netherlands				
	LMG 2749	Human	Unknown				
Pantoea citrea	LMG 22049^T = SHS 2003	Mandarin orange	Japan	6	5	14	
Pantoea punctata	LMG 22050^T = SHS 2006	Mandarin orange	Japan	8	6	21	
	LMG 22097 = SHS 2004	Mandarin orange	Japan				
	LMG 22098 = SHS 2005	Persimmon	Japan				
	LMG 23562 = SHS 2004	Mandarin orange	Japan				
	LMG 23563 = SHS 2007	Mandarin orange	Japan				
	CCUG 30157 = SHS 2004	Mandarin orange	Japan				
	CCUG 30160 = SHS 2007	Mandarin orange	Japan				
Pantoea terrea	LMG 22051^T = SHS 2008	Soil	Japan	8	6	66	
	LMG 23564 = SHS 2009	Soil	Japan				
	LMG 23565 = SHS 2010	Soil	Japan			87	
	CCUG 30162 = SHS 2009	Soil	Japan				
	CCUG 30163 = SHS 2010	Soil	Japan				
Pantoea sp. (MLSA group A)	LMG 24199^R = R-21566	*Eucalyptus*	Uganda	65	20		100
	LMG 24195 = R-24584	*Eucalyptus*	Uruguay				99
	LMG 24196 = R-25674	*Eucalyptus*	Argentina	67			89
	LMG 24201 = R-30991	Maize	South Africa	63			89
	BCC 002	*Eucalyptus*	Argentina				
	BCC 004	*Eucalyptus*	Argentina				
	BCC 006	*Eucalyptus*	Argentina				
	BCC 067	*Eucalyptus*	Colombia				
	BCC 072	*Eucalyptus*	Uruguay	60			94
	BCC 075	*Eucalyptus*	Uruguay				
	BCC 079	*Eucalyptus*	Uruguay				
	BCC 081	*Eucalyptus*	Uruguay	63			99
	BCC 082	*Eucalyptus*	Uruguay				
	BCC 107	*Eucalyptus*	Uganda				
	BCC 208	*Eucalyptus*	Uganda	63			98
	BCC 427	*Eucalyptus*	Uganda				
	BCC 756	*Eucalyptus*	Uruguay				
	BCC 757	*Eucalyptus*	Uruguay				
	BCC 760	*Eucalyptus*	Uruguay				
	BD 502	Maize	South Africa				
Pantoea sp. (MLSA group B)	LMG 24197^R = R-25678	*Eucalyptus*	Uruguay	56	21		100
	LMG 24198 = R-25679	*Eucalyptus*	Uruguay	60			99

Please cite this article as: C. Brady, et al., Phylogeny and identification of *Pantoea* species associated with plants, humans and the natural environment based on multilocus sequence analysis..., Syst. Appl. Microbiol. (2008), doi:10.1016/j.syapm.2008.09.004

Pantoea sp. (MLSA group C)	LMG 2558^{R} = NCPPB 1682	Balsam	India	56	29		100
	LMG 2560 = NCPPB 1941	Marigold	Unknown				105
Pantoea sp. (MLSA group D)	LMG 24200^{R} = R-31523	*Eucalyptus*	Uganda	54	29		100
Pantoea sp. (MLSA group E)	LMG 5343^{R} = CDC 3482-71	Human	USA	48	29		100
	LMG 24532 = CDC 2928-68	Human	USA				
	LMG 24533 = CDC 2525-70	Human	USA				
Pantoea sp. (MLSA group F)	LMG 24534^{R} = CDC 3527-71	Human	USA				100
Pantoea sp. (MLSA group G)	LMG 24202 = R-27853	Onion	USA		57		100
	LMG 24203 = R-21588	Onion	South Africa		44		
	LMG 24248^{R} = R-27856	Onion	South Africa	26	55		
Pantoea sp. (MLSA group H)	LMG 5345^{R} = CDC 3123-70	Human	USA	29	22		100
	LMG 24526 = CDC 238-70	Human	USA				
	LMG 24527 = CDC 1778-70	Human	USA				
	LMG 24528 = CDC 217-71	Human	USA				
Pantoea sp. (MLSA group I)	LMG 2781^{R} = CDC 1741-71	Human	USA	24	19		100
	LMG 5346 [syn. LMG 2781]	Human	USA				
	LMG 24529 = CDC 3638-70	Human	USA				
	LMG 24530 = CDC 5795-70	Human	USA				
	LMG 24531 = CDC 6148-70	Human	USA				
Pantoea citrea (MLSA group J)	LMG 23360^{R}	Pineapple	Philippines			15	100
	LMG 23359	Pineapple	Philippines				
Pantoea sp.	LMG 24194 = R-25665	*Eucalyptus*	Argentina	21	16		
Erwinia billingiae	LMG 2613^{T}	Pear	UK				
Erwinia rhapontici	LMG 2688^{T}	Rhubarb	UK				
Erwinia psidii	LMG 7034^{T}	Guava	Brazil				
Erwinia toletana	LMG 24162^{T}	Olive tree	Spain				
Pectobacterium cypripedii	LMG 2657^{T}	Orchid	USA				
	LMG 2654	Orchid	USA				
	LMG 2655 [syn. LMG 2657]	Orchid	USA				
	LMG 2658	Orchid	Unknown				
Tatumella ptyseos	LMG 7888^{T}	Human	USA			100	

Footnote: LMG = BCCM/LMG Bacteria Collection, Ghent University; BCC = Bacterial Culture Collection, Forestry and Agricultural Biotechnology Institute, Pretoria, South Africa; ATCC = American Type Culture Collection, Rockville, MD, USA; CCUG = Culture Collection, University of Göteborg, Sweden; CDC = Centres for Disease Control, Atlanta, Georgia, USA; BD = Plant Pathogenic and Plant Protecting Bacteria (PPPPB) Culture Collection ARC-PPRI, Pretoria, South Africa; NCPPB = National Collection of Plant Pathogenic Bacteria, York, UK; ^{T} = type strain; ^{R} = reference strain.

[a]Strains hybridized to the type strain of *P. agglomerans* LMG 1286^{T}.

[b]Strains hybridized to the type strain of *P. ananatis* LMG 2665^{T}.

[c]Strains hybridized to the type strain of *T. ptyseos* LMG 7888^{T}.

[d]Strains hybridized to the reference strains selected from MLSA species clusters.

Please cite this article as: C. Brady, et al., Phylogeny and identification of *Pantoea* species associated with plants, humans and the natural environment based on multilocus sequence analysis..., Syst. Appl. Microbiol. (2008), doi:10.1016/j.syapm.2008.09.004

Table 2. Amplification and sequencing primers for *gyrB*, *rpoB*, *atpD* and *infB*

	Sequence [5′→3′]
Amplification primers	
gyrB 01-F	TAA RTT YGA YGA YAA CTC YTA YAA AGT
gyrB 02-R	CMC CYT CCA CCA RGT AMA GTT
rpoB CM7-F	AAC CAG TTC CGC GTT GGC CTG
rpoB CM31b-R	CCT GAA CAA CAC GCT CGG A
atpD 01-F	RTA ATY GGM GCS GTR GTN GAY GT
atpD 02-R	TCA TCC GCM GGW ACR TAW AYN GCC TG
infB 01-F	ATY ATG GGH CAY GTH GAY CA
infB 02-R	ACK GAG TAR TAA CGC AGA TCC A
Sequencing primers	
gyrB 07-F	GTV CGT TTC TGG CCV AG
gyrB 08-R	CTT TAC GRC GKG TCA TWT CAC
rpoB CM81-F	CAG TTC CGC GTT GGC CTG
rpoB CM81b-F	TGA TCA ACG CCA AGC C
rpoB CM32b-R	CGG ACC GGC CTG ACG TTG CAT
atpD 03-F	TGC TGG AAG TKC AGC ARC AG
atpD 04-R	CCM AGY ART GCG GAT ACT TC
infB 03-F	ACG GBA TGA TYA CST TCC TGG
infB 04-R	AGY TTA GAT TTC TGC TGA CG

used in this study. The sequences for both the amplification and sequencing primers for all four genes are listed in Table 2.

PCR and sequencing

Complete 16S rRNA sequences were determined for the type strains of selected members of the *Enterobacteriaceae* using the primers and conditions determined by Coenye et al. [6]. MLSA PCR was performed on each of the strains listed in Table 1 using each of the four sets of amplification primers. Each 50 µl PCR reaction consisted of 5 µl 10× PCR buffer, 5 µl dNTP's (200 µM each), 0.5 µl forward primer (50 µM), 0.5 µl reverse primer (50 µM), 1 µL AmpliTaq DNA polymerase (1 U/µl), 5 µl template DNA and 33 µl sterile MilliQ water. The amplification conditions included denaturation at 95 °C for 5 min, 3 cycles of denaturation at 95 °C for 1 min, annealing at 55 °C for 2 min 15 s and elongation at 72 °C for 1 min 15 s, followed by 30 cycles of denaturation at 95 °C for 35 s, annealing at 55 °C for 1 min 15 s and elongation at 72 °C for 1 min 15 s and a further 7 min of elongation at 72 °C. An annealing temperature of 50 °C was used for several strains which would not amplify at 55 °C. PCR products were separated on 1% agarose gels at 75 V for 45 min. Those reactions resulting in positive PCR products of the expected size were purified using NucleoFast 96 PCR plates (Machery-Nagel). Sequencing reactions were performed using 3 µl purified PCR product, 2 µl 5× sequencing buffer, 0.2 µl Big Dye sequencing reaction mix, 3 µl primer (4 µM) and 1.8 µl sterile MilliQ water. The sequencing conditions included denaturation at 96 °C for 5 s, 25 cycles of denaturation at 96 °C for 10 s, annealing at 55 °C for 10 s and elongation at 60 °C for 4 min. Sequencing reactions were purified using the Montage Seq 96 Sequencing Reaction Cleanup Kit (Millipore). All PCR set-up and purification steps were carried out on a Genesis Workstation 200 (Tecan).

Sequence analysis

Consensus sequences for each strain were assembled by manual alignment of the internal sequences using BioEdit Sequence Alignment Editor v 5.0.9 [21]. Following sequence alignment and trimming of the overhangs using ClustalX [42], the lengths of the four genes were: *gyrB* = 742 bp, *rpoB* = 637 bp, *atpD* = 657 bp and *infB* = 615 bp. A partition-homogeneity test was performed in PAUP 4.0b10 [41] to ensure the four genes could be combined to form a single concatenated data set [10]. Peptide sequences were determined for each of the four genes using BioEdit and were edited manually to ensure no errors were made from alignment gaps. The lengths of the peptide sequences for *gyrB*, *rpoB*, *atpD* and *infB* were 247, 212, 219 and 205 amino acids, respectively. The Modeltest 3.7 programme [34] was then applied to the 16S rRNA data, all four individual gene data sets, as well as the concatenated data set, to determine the best-fit evolutionary model to apply to each gene. Maximum likelihood and neighbour joining trees were drawn using

Please cite this article as: C. Brady, et al., Phylogeny and identification of *Pantoea* species associated with plants, humans and the natural environment based on multilocus sequence analysis..., Syst. Appl. Microbiol. (2008), doi:10.1016/j.syapm.2008.09.004

Phyml [20] and PAUP 4.0b10 [41], respectively by applying the models and parameters determined by Modeltest: the general time reversible model for 16SrRNA, *gyrB*, *rpoB*, *atpD* and the concatenated data set and the Tamura-Nei (TN93) model for *infB*. The individual- and concatenated-peptide sequence trees were drawn in Phyml using the Jones, Taylor and Thornton model and by estimating the proportion of invariable sites and the gamma distribution parameters.

Bootstrap analysis with 1000 replicates was performed on all trees to assess the reliability of the clusters generated. *Escherichia coli*, *Shigella dysenteriae* and *Citrobacter rodentium* were chosen as outgroups and several *Erwinia* species, *Tatumella ptyseos* and *Pectobacterium cypripedii*, the closest phylogenetically related neighbours of *Pantoea* were also included in the MLSA trees [38,46]. The sequences for the four housekeeping genes of *E. coli*, *S. dysenteriae*, *C. rodentium* and *Erwinia amylovora* were obtained from the genome sequencing databases of the Sanger Institute (http://www.sanger.ac.uk) and the University of Wisconsin (https://asap.ahabs.wisc.edu/asap). The genes for *Erwinia billingiae* (LMG 2613^T), *Erwinia rhapontici* (LMG 2688^T), *Erwinia toletana* (LMG 24162), *Erwinia psidii* (LMG 7034), *T. ptyseos* (LMG 7888^T) and *Pe. cypripedii* (LMG 2657^T, LMG 2654, LMG 2655, LMG 2658) were sequenced along with the *Pantoea* strains. The likelihood of alternative tree topologies was evaluated by the Approximately Unbiased (AU) and Shimodaira–Hasegawa (SH) tests using CONSEL [37]. The MLSA data from the four housekeeping genes were compared amongst each other, and to DNA–DNA hybridization values in Bionumerics (Applied Maths), by calculating the correlation between the experiment types, to determine the congruence. A scatter plot was constructed depicting the correlation between *gyrB* sequence similarity (the most congruent), and corresponding DNA–DNA hybridization values.

DNA–DNA hybridization

High-quality DNA for DNA–DNA hybridization of strains was prepared by the method of Wilson [45], with minor modifications [5]. DNA–DNA hybridizations were performed using the microplate method [14] with some modifications [5]. The hybridization temperature was $45 \pm 1\,^{\circ}C$ and reciprocal reactions were performed with DNA from all strains. Representative strains from all novel species clusters were selected and hybridized to the type strains of *P. agglomerans* (LMG 1286^T), *P. ananatis* (LMG 2665^T) and amongst each other.

Results and discussion

Pantoea strains are isolated from environmental sources on a regular basis. The isolates can be human and clinical strains, the causal agents of diseases on plants, epi- and endophytes or merely present in water and soil samples. The ecology and host interactions of these isolates are not clearly understood, often due to the increasing number of strains which are rarely conclusively identified. Therefore, a rapid technique is required to classify and identify these phenotypically related strains to the species level. The difficulties experienced in identifying *Pantoea* species are exacerbated by the uncertain phylogeny, based on 16S rRNA sequences, of the genus.

The 16S rRNA phylogenetic tree (Fig. 1) revealed the polyphyletic nature of the genus *Pantoea*. The type strains of the "core" *Pantoea* species clustered together with reference strains from seven potential novel species of *Pantoea*. The type strain of *P. dispersa* clustered on the border of the genus *Erwinia* with the reference strain of Brenner's DNA group IV, LMG 2781^R [3]. It has been suggested that *Pe. cypripedii* is closely related to *Pantoea* and belongs to that genus [46]. However, in the 16S rRNA phylogeny the type strain of *Pe. cypripedii* grouped on the border of *P. dispersa*, far removed from the *Pantoea* "core" species. The "Japanese" species formed a separate well-supported cluster which included the type strain of *T. ptyseos*, a single species genus of clinical strains. A study by Paradis et al. [33] highlighted a clear phylogenetic affiliation between the genera *Pantoea* and *Tatumella*. Based on that observation, the type strain of *T. ptyseos* was included in our MLSA study. Although it is likely that the "Japanese" species do not actually belong to *Pantoea*, *P. dispersa* is considered a "core" species and its position within the genus must be determined. The polyphyletic nature of *Pantoea* is not unusual as several other genera within the *Enterobacteriaceae* have also been shown to be polyphyletic based on 16S rRNA phylogeny, including *Enterobacter*, *Brenneria* [24], *Erwinia* and *Pectobacterium* [46].

One of the major objectives of this study was to resolve the phylogenetic position of *Pantoea* within the family *Enterobacteriaceae*. It is apparent that this is not possible based on 16S rRNA sequences. However, MLSA data showed great promise for the phylogenetic analysis of the genus *Pantoea* and closely related enterobacteria. A concatenated tree, based on the nucleotide sequences of the four housekeeping genes, appeared to be most reliable for determining phylogenetic relationships amongst *Pantoea* strains (Fig. 2). In this tree, *Pantoea* strains were divided into two clusters, both supported by high bootstrap values of 97% and 100%, respectively. The first cluster contained

Please cite this article as: C. Brady, et al., Phylogeny and identification of *Pantoea* species associated with plants, humans and the natural environment based on multilocus sequence analysis..., Syst. Appl. Microbiol. (2008), doi:10.1016/j.syapm.2008.09.004

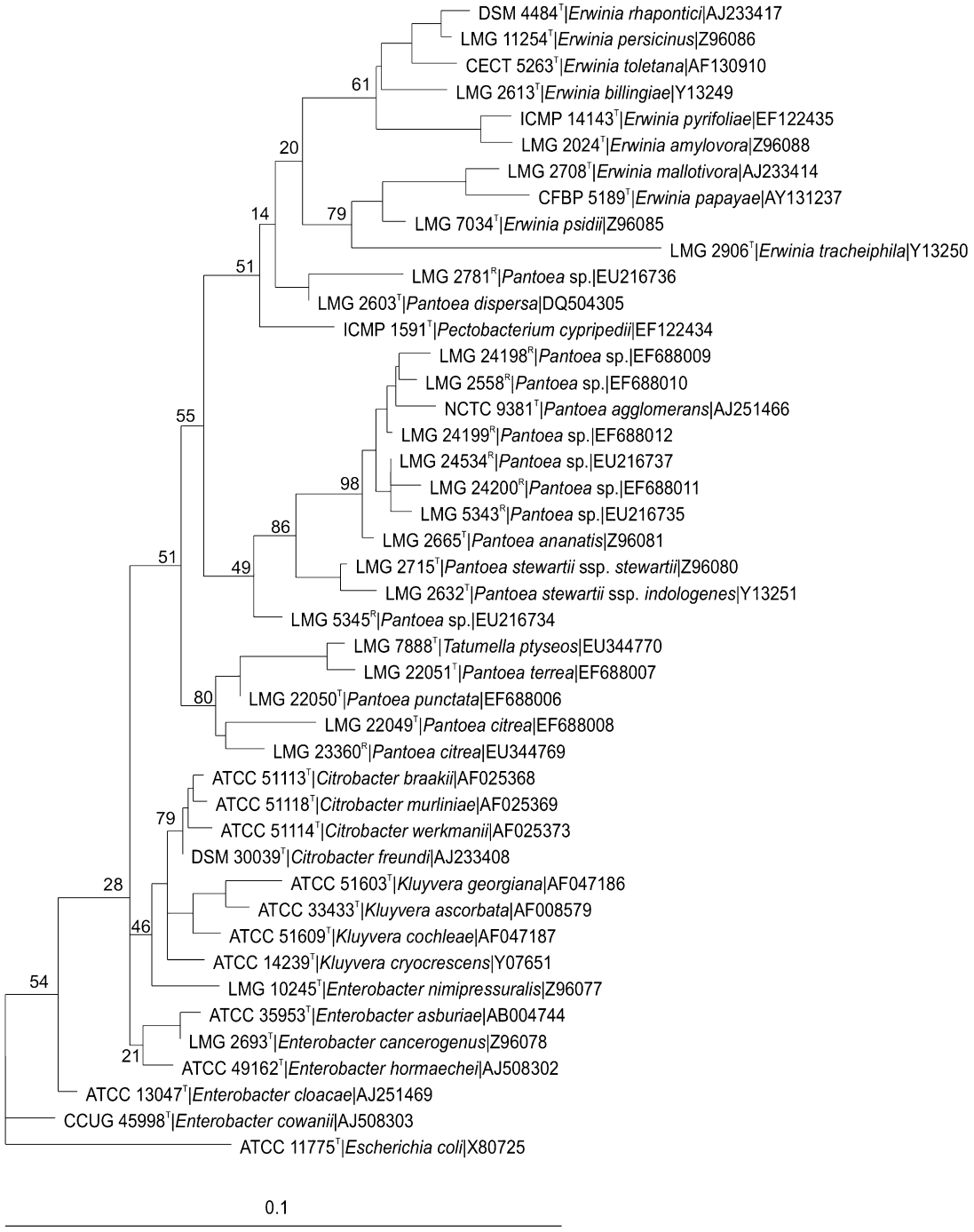

Fig. 1. Maximum likelihood tree based on the 16S rRNA sequences of *Pantoea* species and their closest phylogenetic neighbours. Bootstrap values after 1000 replicates are expressed as percentages. *Escherichia coli* was included as an outgroup.

P. agglomerans, *P. ananatis*, *P. stewartii*, *P. dispersa* (the "core" species), nine MLSA groups of potential novel *Pantoea* species and strains of *Pe. cypripedii*. The second cluster contained the "Japanese species" (*P. punctata*, *P. citrea* and *P. terrea*), the type strain of *T. ptyseos* and a tenth potential novel species. Despite the aberrant position of the type strain of *P. dispersa* in the 16S rRNA tree, it appears that this species is part of the "core" of *Pantoea* as observed in the concatenated tree with high bootstrap support. Additionally *Pe. cypripedii* appears to belong to the genus *Pantoea*, as strains of this species clustered within the "core" also with high bootstrap support. The large distance between the *Pantoea* "core" and "Japanese" species was consistently observed within the concatenated tree and all four individual gene trees (Supplementary data),

Please cite this article as: C. Brady, et al., Phylogeny and identification of *Pantoea* species associated with plants, humans and the natural environment based on multilocus sequence analysis..., Syst. Appl. Microbiol. (2008), doi:10.1016/j.syapm.2008.09.004

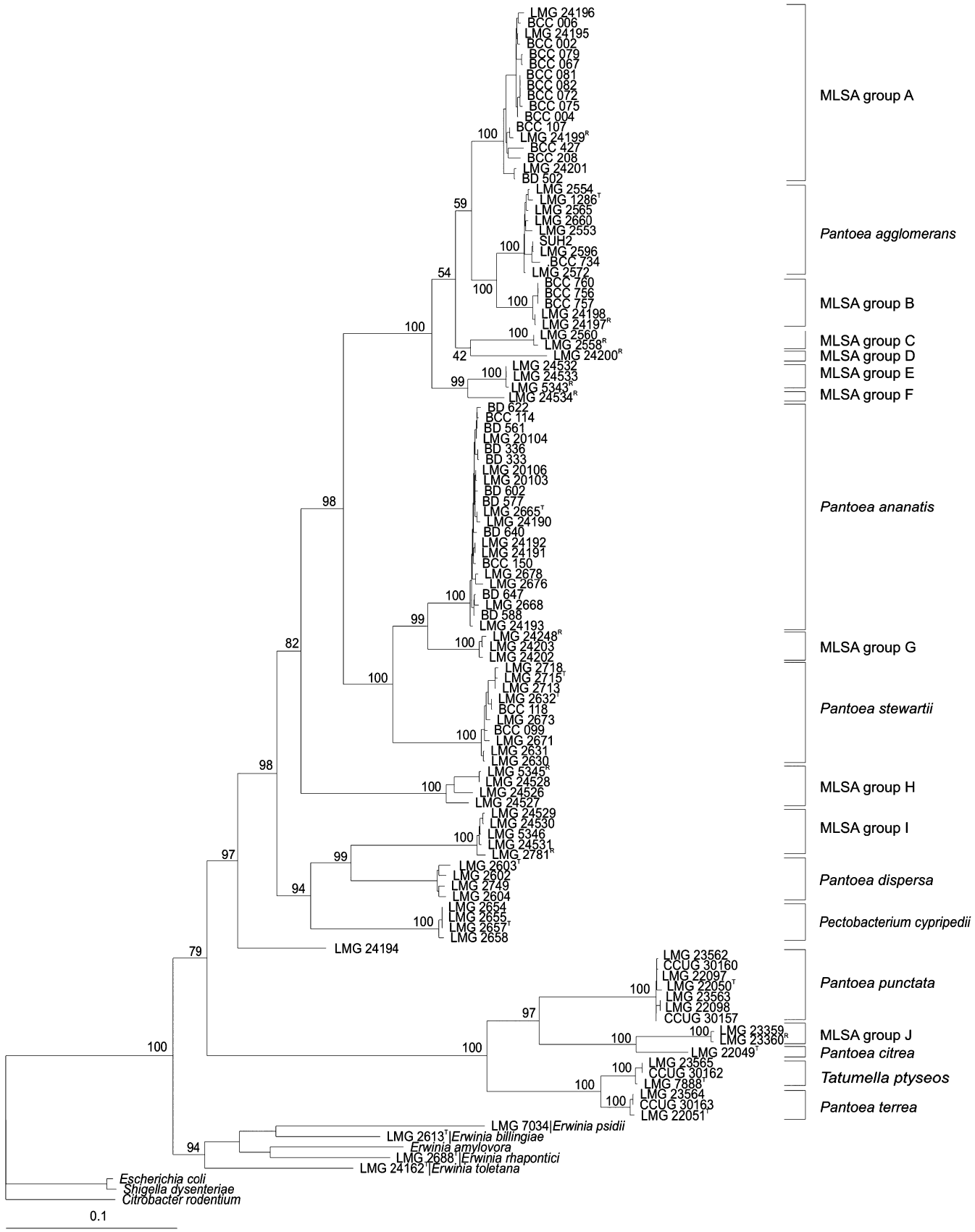

Fig. 2. Maximum likelihood tree based on the concatenated nucleotide sequences of *gyrB*, *rpoB*, *atpD* and *infB* of 103 *Pantoea* strains. Bootstrap values after 1000 replicates are expressed as percentages. *Citrobacter rodentium* was included as an outgroup.

Please cite this article as: C. Brady, et al., Phylogeny and identification of *Pantoea* species associated with plants, humans and the natural environment based on multilocus sequence analysis..., Syst. Appl. Microbiol. (2008), doi:10.1016/j.syapm.2008.09.004

casting serious doubt on the inclusion of these species within the genus and supporting the statement of Grimont and Grimont [19]. In the concatenated tree and all individual gene trees, the "Japanese" species clustered with *T. ptyseos* with bootstrap support of 95–100%, and should be considered to belong to the genus *Tatumella*.

The phylogeny of the concatenated-peptide sequence tree (Suppl. Fig. S1) is in agreement with that of the concatenated nucleotide sequence tree (Fig. 2). *Pe. cypripedii* retained its position in the peptide sequence tree clustering closely with the *P. dispersa* group with a bootstrap of 89%, giving further support for the inclusion of this species within the genus *Pantoea*. The "core" *Pantoea* species were contained in one large cluster, clearly separated from the "Japanese" species and *T. ptyseos* by extended branch lengths and high bootstrap support. Additionally, the DNA–DNA hybridization values support the phylogenetic split of the "Japanese" species from the *Pantoea* "core" species (Table 1). When the type strain of *T. ptyseos* (LMG 7888^T) was hybridized to the type strains of *P. citrea* (LMG 22049^T), *P. punctata* (LMG 22050^T) and *P. terrea* (LMG 22051^T), similarity values ranging between 20% and 66% were observed. These values were considerably higher than those observed when the same type strains were hybridized to the type strains of "core" species of *Pantoea*, where similarity values of less than 9% were observed. This lends further support for the transfer of *P. citrea*, *P. punctata* and *P. terrea* to the genus *Tatumella*.

Alternative user-defined tree topologies, where *Pantoea* was defined as either monophyletic or polyphyletic, were tested using CONSEL. The *P*-values for the AU and SH tests (Suppl. Table S1) indicated the rejection of the likelihood that *Pantoea* is a monophyletic genus. The polyphyletic topology (where *P. dispersa*, MLSA group I and *Pe. cypripedii* as well as the "Japanese" species clustered separate from the "core" species) was supported for all five data sets tested, with *P*-values close to 1.000 and was, therefore, not rejected. This supports the observation in the 16S rRNA phylogenetic tree that *Pantoea* is polyphyletic (Fig. 1), although the same cannot be said for the current concatenated tree (Fig. 2) where *P. dispersa*, MLSA group I and *Pe. cypripedii* do cluster with *Pantoea* "core" species, but at a lower level.

A recent study examining the relationships of plant pathogenic enterobacteria based on the housekeeping genes *atpD*, *carA* and *recA*, suggested that there is no justification for the separation of *Erwinia* and *Pantoea* into two separate genera [46]. However in the present study, *Erwinia* species cluster at a lower level to the "Japanese" species and *Tatumella* in the concatenated nucleotide sequence and peptide sequence trees (Fig. 2 and Suppl. Fig. S1) clearly indicating that *Erwinia* and *Pantoea* should not be united into a single genus.

The individual nucleotide sequence trees (Suppl. Figs. S2a–d) have similar groupings to those of the concatenated nucleotide sequence tree, with two exceptions. In the *gyrB* nucleotide sequence tree (Suppl. Fig. S2a) *P. dispersa*, *Pe. cypripedii* and strains belonging to Brenner DNA group IV (MLSA group I) formed a cluster which was separated from the "core" *Pantoea* species by the "Japanese" species and *Erwinia* species. The same partition was observed in the 16S rRNA phylogenetic tree (Fig. 1). *Pe. cypripedii* strains also clustered separately in the *rpoB* tree (Suppl. Fig. S2b), on the border of the "Japanese" species whilst *P. dispersa* and MLSA group I grouped with the "core" *Pantoea* species. The reasons for these unusual groupings are unclear; however, it does support the preference of making phylogenetic inferences based on concatenated sequence trees rather than individual gene trees [11,25].

Another cluster with different positions in the nucleotide sequence trees contains the single strain LMG 24194. In the concatenated tree, *atpD* and *infB* trees, this strain grouped on the border of, or within, the "core" *Pantoea* clade, but in the *gyrB* and *rpoB* trees LMG 24194 grouped with *Erwinia* species (Fig. 2, Suppl. Figs. S2a–d). LMG 24194 was thought to belong to a potential new *Pantoea* species but as this strain does not retain its position in all five trees, its taxonomic position cannot be clearly concluded at present.

Together and separately all four housekeeping genes, *gyrB*, *rpoB*, *atpD* and *infB*, could delineate the seven validly published *Pantoea* species and revealed ten potential novel species (MLSA groups A–J) in each of the maximum likelihood trees constructed based on nucleotide sequences (neighbour joining trees not shown). The majority of the validly published *Pantoea* species and MLSA groups A–J were supported by strong bootstrap values. The only exceptions included marginally lower intra- and inter-species bootstraps in the *rpoB* tree (Suppl. Fig. S2b) and several weaker inter-species bootstrap values in the *atpD* (Suppl. Fig. S2c) and *infB* (Suppl. Fig. S2d) trees. The bootstrap values for the concatenated data set tree were by far the highest and most stable, not only at the intra-species level, but also between species of the genus (Fig. 2). Strains of the same *Pantoea* species (including type and reference strains) had at least 96.4% *gyrB*, 98.9% *rpoB*, 98.3% *atpD* and 97.2% *infB* gene sequence similarity, whereas at the inter-species level the sequence similarity was at a maximum of 94.8% for *gyrB*, 98.4% for *rpoB*, 97.9% for *atpD* and 96.9% for *infB*.

The largest MLSA group included isolates from *Eucalyptus* showing symptoms of bacterial blight in Uganda, Argentina and Uruguay as well as strains isolated from maize infected with brown stalk rot in

Please cite this article as: C. Brady, et al., Phylogeny and identification of *Pantoea* species associated with plants, humans and the natural environment based on multilocus sequence analysis..., Syst. Appl. Microbiol. (2008), doi:10.1016/j.syapm.2008.09.004

South Africa. In all of the nucleotide sequence trees (Fig. 2, Suppl. Figs. S2a–d), this MLSA group A clustered in close proximity to *P. agglomerans* strains, but distinctly diverged into a separate cluster. MLSA group B also clustered close to *P. agglomerans* and contained strains from *Eucalyptus* infected with bacterial blight, but only from trees in Uruguay. Two strains from a study by Beji et al. [1], LMG 2558^{R} and LMG 2560, were assigned to *P. agglomerans* based on their protein profiles but later excluded from the species by Gavini et al. [17]. Based on the MLSA data presented in this study these two strains, forming MLSA group C, constitute a potential novel species. A single strain, LMG 24200^{R}, formed MLSA group D. This strain was isolated from infected *Eucalyptus* in Uganda and was expected to cluster in MLSA group A along with LMG 24199^{R} and BCC 107, also isolated from Uganda. However, as this strain retained its position in all five trees, LMG 24200^{R} also represents a potential new species of *Pantoea*. MLSA group E contained human strains belonging to the group referred to as Brenner DNA group V, whilst a single strain from the same group (LMG 24534^{R}) constituted MLSA group F. The isolates from onion formed MLSA group G, which consistently branched off from the *P. ananatis* cluster with strong bootstrap support in the majority of the trees. MLSA groups H and I were comprised of human strains from Brenner DNA groups II and IV, respectively. MLSA groups A–I all group within the first subcluster, which is referred to as the "core" *Pantoea* group with *P. agglomerans*, *P. ananatis*, *P. stewartii* and *P. dispersa* [19]. MLSA group J fell within the second subcluster, the "Japanese" *Pantoea* clade, which joined the "core" *Pantoea* clade at a lower level in all of the phylogenetic trees. MLSA group J contained two strains thought to be *P. citrea* as they were identified as the causal agent of pink disease of pineapple [35]. However, MLSA group J formed a cluster distinct from the type strain of *P. citrea* (LMG 22049^{T}) with an extended branch length that is observed in all five trees. Two strains classified as *P. terrea* (LMG 23565 and CCUG 30163) clustered with the type strain of *T. ptyseos* (LMG 7888^{T}) instead of with the type strain of *P. terrea* (LMG 22051^{T}), suggesting that they were wrongly classified.

The potential species clusters observed in this MLSA study were confirmed by DNA–DNA hybridization data (Table 1). Reference strains were selected from each MLSA group, and when hybridized to other strains within that group, the DNA similarity values ranged from 83% to 105%. Representative strains were selected from MLSA groups A–I for hybridization with type strains of *P. agglomerans* and *P. ananatis*. The DNA similarity values were between 24% and 67% for strains hybridized to *P. agglomerans*, and between 19% and 57% for strains hybridized to *P. ananatis*. These data support the inclusion of MLSA groups A–I in the genus *Pantoea* as novel species. When the reference strain of MLSA group J was hybridized to the type strain of *P. citrea*, the DNA similarity value was 42%, supporting the separation of MLSA group J (thought to belong to *P. citrea*) from this species. The DNA similarity values between the reference strain of MLSA group J and the type strains of *P. terrea*, *P. punctata* and *T. ptyseos* were all less than 18%. It is likely that MLSA group J is a novel species belonging to the genus *Tatumella*.

The congruence between the four genes used in this study was determined and varied from 86.3% to 92.6%. The congruence between MLSA and DNA–DNA hybridization data (data not shown) ranged from 84.5% to 95.2%, with *gyrB* being most congruent (95.2%) and *atpD* the least (84.5%). A scatter plot comparing the DNA–DNA hybridization values and *gyrB* sequence data is presented in Fig. 3. It was generally observed that strains sharing more than 70% DNA similarity have high *gyrB* sequence similarity. The only notable exceptions are strains belonging to the subspecies of *P. stewartii* which share lower DNA relatedness (60–65%) but high *gyrB* sequence similarity ($\pm$99%), indicated by an arrow in Fig. 3. The lower DNA–DNA hybridization values could suggest that the two subspecies of *P. stewartii*, subspecies *stewartii* and subspecies *indologenes*, ought to be divided into separate

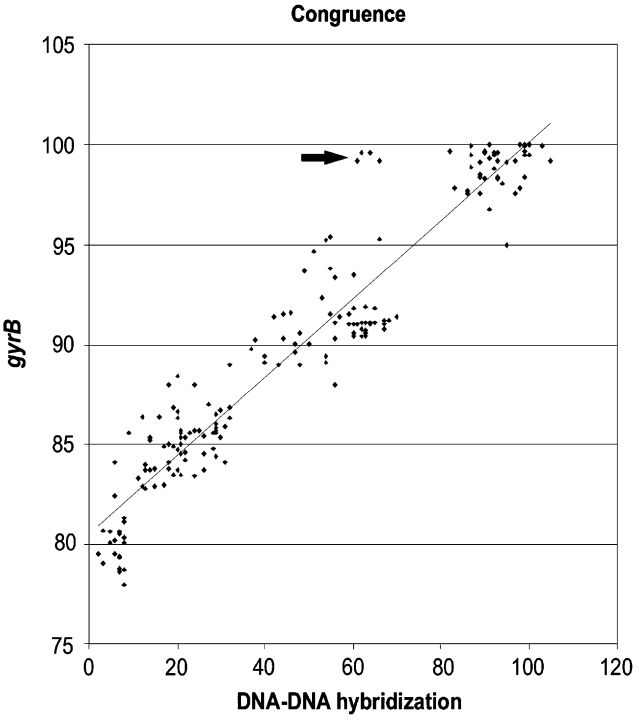

Fig. 3. Scatter plot comparing the congruence of *gyrB* sequence data and DNA–DNA hybridization values for the genus *Pantoea*. The arrow indicates strains of *P. stewartii* which have lower DNA–DNA hybridization values, but high *gyrB* sequence similarity.

Please cite this article as: C. Brady, et al., Phylogeny and identification of *Pantoea* species associated with plants, humans and the natural environment based on multilocus sequence analysis..., Syst. Appl. Microbiol. (2008), doi:10.1016/j.syapm.2008.09.004

species. However, the high *gyrB* sequence similarity supports their status as subspecies of *P. stewartii*.

The MLSA technique was examined for usefulness in the identification and classification of *Pantoea* strains and was found to be successful. The ''Japanese'' *Pantoea* species cluster at a much lower level to the genus and are more closely related to the genus *Tatumella*. Based on the MLSA and DNA–DNA hybridization data presented in this study, the ''Japanese'' species should be transferred to *Tatumella*. *Pe. cypripedii* was shown to be a close phylogenetic relative of *Pantoea* and it is highly probable that this species should belong to *Pantoea*. The high bootstrap values observed at the species level in all phylogenetic trees indicate that the four housekeeping genes used in this study, *gyrB*, *rpoB*, *atpD* and *infB*, are reliable genetic markers for differentiation of *Pantoea* species. Not only could the MLSA scheme distinguish between the seven validly published species of *Pantoea*, it also revealed ten potential new species. The potential species observed in each of the MLSA trees in this study are strongly supported by both AFLP analysis [2] and DNA–DNA hybridization data. The potential new *Pantoea* species are now in the process of being described using the MLSA data as a supporting technique. In conclusion, MLSA provides a rapid technique for reliable identification and classification of *Pantoea* strains to the species level and is clearly more discriminatory than 16S rRNA sequencing.

Acknowledgements

This study was funded by the South African–Flemish Bilateral Agreement, the National Research Foundation (NRF), the Tree Protection Co-operative Programme (TPCP) and the THRIP support programme of the Department of Trade and Industry, South Africa. The BCCM/LMG Bacteria collection is supported by the Federal Public Planning Service-Science Policy, Belgium. The authors wish to acknowledge Bart Hoste, Katrien Vandemeulebroecke, Dirk Gevers and Peter Dawyndt for their help with the primer design and technical assistance. We thank Mike Wingfield for collecting the diseased *Eucalyptus* material, Teresa Goszczynska for allowing us to include the strains isolated from onion and maize, Mrs. Mohr for providing the Brenner strains from the CDC, and Emma Steenkamp for helping with the phylogenetic analysis.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.syapm.2008.09.004.

References

[1] A. Beji, J. Mergaert, F. Gavini, D. Izard, K. Kersters, H. Leclerc, J. De Ley, Subjective synonymy of *Erwinia herbicola*, *Erwinia milletiae*, and *Enterobacter agglomerans* and redefinition of the taxon by genotypic and phenotypic data, Int. J. Syst. Bacteriol. 38 (1988) 77–88.

[2] C. Brady, S. Venter, I. Cleenwerck, M. Vancanneyt, J. Swings, T. Coutinho, A FAFLP system for the improved identification of plant-pathogenic and plant-associated species of the genus *Pantoea* 30 (2007) 413–417.

[3] D.J. Brenner, G.R. Fanning, J.K.L. Knutson, A.G. Steigerwalt, M.I. Krichevsky, Attempts to classify Herbicola group-*Enterobacter agglomerans* strains by deoxyribonucleic acid hydridization and phenotypic tests, Int. J. Syst. Bacteriol. 34 (1984) 45–55.

[4] T.J. Burr, B.H. Katz, G.S. Abawi, D.C. Crosier, Comparison of tumorigenic strains of *Erwinia herbicola* isolated from table beet with *E.h. gypsophilae*, Plant Dis. 75 (1991) 855–858.

[5] I. Cleenwerck, K. Vandemeulebroecke, D. Janssens, J. Swings, Re-examination of the genus *Acetobacter*, with descriptions of *Acetobacter cerevisiae* sp. nov. and *Acetobacter malorum* sp. nov., Int. J. Syst. Evol. Microbiol. 52 (2002) 1551–1558.

[6] T. Coenye, E. Falsen, M. Vancanneyt, B. Hoste, J.R.W. Govan, K. Kersters, P. Vandamme, Classification of *Alcaligenes faecalis*-like isolates from the environment and human clinical samples as *Ralstonia gilardii* sp. nov., Int. J. Syst. Bacteriol. 49 (1999) 405–413.

[7] D.A. Cooksey, Galls of *Gypsophila paniculata* caused by *Erwinia herbicola*, Plant Dis. 70 (1986) 464–468.

[8] E.J. Cother, R. Reinke, C. McKenzie, V.M. Lanoiselet, D.H. Noble, An unusual stem necrosis of rice caused by *Pantoea ananas* and the first record of this pathogen on rice in Australia, Austr. Plant Pathol. 33 (2004) 495–503.

[9] T.A. Coutinho, O. Preisig, J. Mergaert, M.C. Cnockaert, K.H. Riedel, J. Swings, M.J. Wingfield, Bacterial blight and dieback of *Eucalyptus* species, hybrids, and clones in South Africa, Plant Dis. 86 (2002) 20–25.

[10] C.W. Cunningham, Can three incongruence tests predict when data should be combined?, Mol. Biol. Evol. 14 (1997) 733–740.

[11] C. Dauga, Evolution of the *gyrB* gene the molecular phylogeny of *Enterobacteriaceae*: a model molecule for molecular systematic studies, Int. J. Syst. Evol. Microbiol. 52 (2002) 531–547.

[12] T. De Baere, R. Verhelst, C. Labit, G. Verschraegen, G. Wauters, G. Claeys, M. Vaneechoutte, Bacteremic infection with *Pantoea ananatis*, J. Clin. Microbiol. 42 (2004) 4393–4395.

[13] M. Drancourt, C. Bollet, A. Carta, P. Rousselier, Phylogenetic analyses of *Klebsiella* species delineate *Klebsiella* and *Raoultella* gen. nov., with description of *Raoultella ornithinolytica* comb. nov., *Raoultella terrigena* comb. nov. and *Raoultella planticola* comb. nov., Int. J. Syst. Evol. Microbiol. 51 (2001) 925–932.

[14] T. Ezaki, Y. Hashimoto, E. Yabuuchi, Fluorometric deoxyribonucleic acid-deoxyribonucleic acid hybridization in micro-dilution wells as an alternative to membrane filter hybridization in which radioisotopes are used to

determine genetic relatedness among bacterial strains, Int. J. Syst. Bacteriol. 39 (1989) 224–229.
[15] D.G. Fullerton, A.A. Lwin, S. Lal, *Pantoea agglomerans* liver abscess presenting with a painful thigh, Eur. J. Gastroenterol. Hepatol. 19 (2007) 433–435.
[16] F. Gavini, B. Lefebvre, H. Leclerc, Taxonomic study of strains belonging or related to the genus *Erwinia*, *herbicola* group, and to the species *Enterobacter agglomerans*, Syst. Appl. Microbiol. 4 (1983) 218–235.
[17] F. Gavini, J. Mergaert, A. Beji, C. Mielcarek, D. Izard, K. Kersters, J. de Ley, Transfer of *Enterobacter agglomerans* (Beijerinck 1888) 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 (1989) 337–345.
[18] T. Goszczynska, W.J. Botha, S.N. Venter, T.A. Coutinho, Isolation and identification of the causal agent of brown stalk rot, a new disease of maize in South Africa, Plant Dis. 91 (2007) 711–718.
[19] P.A.D. Grimont, F. Grimont, Genus: Pantoea, in: D.J. Brenner, N.R. Krieg, J.T. Staley (Eds.), Bergey's Manual of Systematic Bacteriology, vol. 2, The *Proteobacteria*, Part B, The *Gammaproteobacteria*, second ed., Springer, New York, 2005, pp. 713–720.
[20] S. Guidon, O. Gascual, A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood, Syst. Biol. 52 (2003) 696–704.
[21] T.A. Hall, BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, Nucl. Acids Symp. Ser. 41 (1999) 95–98.
[22] J. Hedegaard, S.A. Steffensen, N. Norskov-Lauritsen, K.K. Mortensen, H.U. Sperling-Petersen, Identification of *Enterobacteriaceae* by partial sequencing of the gene encoding translation initiation factor 2, Int. J. Syst. Bacteriol. 49 (1999) 1531–1538.
[23] B. Kageyama, M. Nakae, S. Yagi, T. Sonoyama, *Pantoea punctata* sp. nov., *Pantoea citrea* sp. nov., and *Pantoea terrea* sp. nov. isolated from fruit and soil samples, Int. J. Syst. Bacteriol. 42 (1992) 203–210.
[24] P. Kämpfer, S. Ruppel, R. Remus, *Enterobacter radicincitans* sp. nov., a plant growth promoting species of the family *Enterobacteriaceae*, Syst. Appl. Microbiol. 28 (2005) 213–221.
[25] K.T. Konstantinidis, A. Ramette, J.M. Tiedje, Toward a more robust assessment of intraspecies diversity, using fewer genetic markers, Appl. Environ. Microbiol. 72 (2006) 7286–7293.
[26] Kratz, D. Greenberg, Y. Barki, E. Cohen, M. Lifshitz, *Pantoea agglomerans* as a cause of septic arthritis after palm tree thorn injury; case report and literature review, Arch. Dis. Child. 88 (2003) 542–544.
[27] X. Li, D. Zhang, F. Chen, J. Ma, Y. Dong, L. Zhang, *Klebsiella singaporensis* sp. nov., a novel isomaltulose-producing bacterium, Int. J. Syst. Evol. Microbiol. 54 (2004) 2131–2136.
[28] P. Lim, S. Chen, C. Tsai, M. Pai, *Pantoea* peritonitis in a patient receiving chronic ambulatory peritoneal dialysis, Nephrology 11 (2006) 97–99.
[29] G.C. McGhee, E.L. Schnabel, K. Maxson-Stein, B. Jones, V.K. Stromberg, G.H. Lacy, A.L. Jones, Relatedness of chromosomal and plasmid DNAs of *Erwinia pyrifoliae* and *Erwinia amylovora*, Appl. Environ. Microbiol. 68 (2002) 6182–6192.
[30] J. Mergaert, L. Verdonck, K. Kersters, 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 (1993) 162–173.
[31] C. Mollet, M. Drancourt, D. Raoult, *rpoB* Sequence analysis as a novel basis for bacterial identification, Mol. Microbiol. 26 (1997) 1005–1011.
[32] S. Niemann, A. Puehler, H.-V. Tichy, R. Simon, W. Selbitschka, Evaluation of the resolving power of three different DNA fingerprinting methods to discriminate among isolates of a natural *Rhizobium meliloti* population, J. Appl. Microbiol. 82 (1997) 477–484.
[33] S. Paradis, M. Boissinot, N. Paquette, S.D. Belanger, E.A. Martel, D.K. Boudreau, F.J. Picard, M. Ouellette, P.H. Roy, M.G. Bergeron, Phylogeny of the *Enterobacteriaceae* based on genes encoding elongation factor Tu and F-ATPase beta-subunit, Int. J. Syst. Evol. Microbiol. 55 (2005) 2013–2025.
[34] D. Posada, K.A. Crandall, Modeltest: testing the model of DNA substitution, Bioinformatics 14 (1998) 817–818.
[35] C.J. Pujol, C.I. Kado, Genetic and biochemical characterization of the pathway in *Pantoea citrea* leading to pink disease of pineapple, J. Bacteriol. 182 (2000) 2230–2237.
[36] H. Schmid, S. Schubert, C. Weber, J.R. Bogner, Isolation of a *Pantoea dispersa*-like strain from a 71-year-old woman with acute myeloid leukemia and multiple myeloma, Infection 31 (2003) 66–67.
[37] H. Shimodaira, M. Hasegawa, CONSEL: for assessing the confidence of phylogenetic tree selection, Bioinformatics 17 (2001) 1246–1247.
[38] C. Spröer, U. Mendrock, J. Swiderski, E. Lang, E. Stackebrandt, The phylogenetic position of Serratia, Buttiauxella and some other genera of the family Enterobacteriaceae, Int. J. Syst. Bacteriol. 49 (1999) 1433–1438.
[39] R. Stephan, S. Van Trappen, I. Cleenwerck, M. Vancanneyt, P. De Vos, A. Lehner, *Enterobacter turicensis* sp. nov. and *Enterobacter helveticus* sp. nov., isolated from fruit powder, Int. J. Syst. Evol. Microbiol. 57 (2007) 820–826.
[40] F.C. Stewart, A bacterial disease of sweet corn, New York State Agric. Exp. Stn. Bull. 130 (1897) 422–439.
[41] D.L. Swofford, PAUP*: Phylogenetic Analysis Using Parsimony and Other Methods (software), Sinauer Associates, Sunderland, MA, 2000.
[42] J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, D.G. Higgins, The ClustalX-Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, Nucl. Acids Res. 25 (1997) 4876–4882.
[43] H. Van Rostenberghe, R. Noraida, W.I. Wan Pauzi, H. Habsah, M. Zeehaida, A.R. Rosliza, I. Fatimah, N.Y. Nik Sharimah, H. Maimunah, The clinical picture of

Please cite this article as: C. Brady, et al., Phylogeny and identification of *Pantoea* species associated with plants, humans and the natural environment based on multilocus sequence analysis..., Syst. Appl. Microbiol. (2008), doi:10.1016/j.syapm.2008.09.004

neonatal infection with *Pantoea* species, Jpn. J. Infect. Dis. 59 (2006) 120–121.
[44] L. Verdonck, J. Mergaert, C. Rijckaert, J. Swings, K. Kersters, J. De Ley, Genus *Erwinia*: Numerical analysis of phenotypic features, Int. J. Syst. Bacteriol. 37 (1987) 4–18.
[45] K. Wilson, Preparation of genomic DNA from bacteria, in: F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, K. Struhl (Eds.), Current Protocols in Molecular Biology, Green Publishing and Wiley-Interscience, New York, 1987, pp. 2.4.1–2.4.5.
[46] J.M. Young, D.C. Park, Relationships of plant pathogenic enterobacteria based on partial *atpD*, *carA*, and *recA* as individual and concatenated nucleotide and peptide sequences, Syst. Appl. Microbiol. 30 (2007) 343–354.

Please cite this article as: C. Brady, et al., Phylogeny and identification of *Pantoea* species associated with plants, humans and the natural environment based on multilocus sequence analysis..., Syst. Appl. Microbiol. (2008), doi:10.1016/j.syapm.2008.09.004