

Supplementary File 1

Isolate collection

For the genealogical concordance analysis to distinguish between *Pantoea ananatis* and *Pantoea allii* a diverse set of isolates were used. The isolates included in this study are listed in the table below.

Species	Isolate number	Origin
<i>Pantoea ananatis</i>	LMG 2665 ¹	Brazil, Pineapple
	B 1-9	Korea, Rhizosphere
	LMG 5342	Georgia, Human
	SUPP 1791	Japan, Melon
	BCC 147	Japan, Rice
	BCC 154	Hawaii, Pineapple
	BCC 376	Australia, Rice
	BCC 601	Brazil, Sugarcane
	BCC 155	Brazil, <i>Cattleya</i> sp.
	y101	Japan, <i>Artemisia</i> sp.
	BD 664	South Africa, <i>Eucalyptus</i>
	BCC 151	U.S.A, Honey melons
	BCC 104	U.S.A., Sudangrass
	BD 316	South Africa, Onion
	BCC 644	Unknown, Banana
	BCC 149	Hungary, Rust
	BCC 367	Thailand, <i>Eucalyptus</i>
<i>Pantoea allii</i>	BD304	South Africa, Onion
	BD309	U.S.A, Onion
	BD377	South Africa, Onion
	BD381	South Africa, Onion
	BD391	South Africa, Onion
	LMG 24248 ¹	South Africa, Onion
<i>Pantoea stewartii</i> subsp. <i>stewartii</i>	LMG 2715 ¹	U.S.A, Maize
<i>Pantoea stewartii</i> subsp. <i>indologenes</i>	LMG 2632 ¹	India, Fox millet

Gene sequences

Apart from *gyrB*, *infB* and *atpD* (Brady et al. 2011), two additional genes were partially sequenced. They were *ompF* and *pmrB*, encoding for an outer membrane porin precursor F and a sensor protein of two component system that regulates lipopolysaccharides modification. The PCR and sequencing primers were designed based on the available

sequence data and are listed below.

Primer	Sequence (5' → 3')	Length	GC %	Tm (°C)
ompF-F	CGT CTG TGC GAC AGA ATT ACC	21	52	62.5
ompF-R	GAT GAA GCG CAA CAT TCT GGC AG	23	52	64.5
ompF-S	TTA GAC TGA ACG TAG GCC AG	20	50	60.4
pmrB-F	TAC CGT TCT GGC AAT ACC ATG C	22	50	62.6
pmrB-R	TTA GCG TAG CAG GGA TCG	18	55	59.9

The PCR reaction mixture consisted 2.5 µl of 10X Reaction buffer (Separation Scientific), 2 µl of 25mM MgCl₂ (Separation Scientific), 2 µl of 250µM dNTPs (Fermentas), 0.5 µl of each 5pmol forward and reverse primer (Inqaba Biotec), 0.3 µl of 5U/µl Super-therm polymerase (Separation Scientific) and 16.2 µl nuclease free water (WhiteSci), making up a total volume of 25 µl.

The PCR protocol, performed on a Veriti™ Thermal Cycler (Applied Biosystems), included initiation at 94°C for 10 minutes which was followed by 30 cycles of 1 minute denaturation at 92°C, 1 minute annealing at 60°C and 1 minute extension at 72°C, ended with a final extension at 72°C for 7 minutes. The verification of PCR products was done by electrophoresis on an agarose gel.

The sequencing reaction was made up of 4 µl of cleaned PCR product, 2 µl of 5X Sequencing buffer (Applied Biosystems), 1 µl of v3.1 ABI PRISM® BigDye Terminator (Applied Biosystems), 0.3 µl of a 100 µM solution of the specific sequencing primer and 4.7 µl nuclease free water (Qiagen) resulting in a final volume of 12 µl. The following sequencing PCR cycle was run: 5 seconds initiation at 96°C, 25 cycles of denaturation at 96°C for 10 seconds, annealing at 55°C for 5 seconds and 4 minutes elongation at 60°C. The products were sequenced on an ABI PRISM® DNA Automated Sequencer (Perkin-Elmer) at the University of Pretoria's DNA Sequencing Facility.

All of the *Pan. ananatis* and *Pan. allii* sequences generated for this study has been submitted to the European Nucleotide Archive (<http://www.ebi.ac.uk>). Accession numbers for the five genes studied in the *Pantoea* isolates are provided below:

Species	Locus	Isolate name	Accession Number/Locus Tag		
<i>Pan. ananatis</i>	<i>atpD</i>	B1-9	NZ_CAEJ01000140		
		BCC367	LT708142		
		BCC376	LT708137		
		LMG5342	PANA5342_RS22945		
		BCC132 = LMG2665 ^T	EF988737		
		BCC154 = LMG2664	LT708136		
		BCC149 = LMG2675	LT708143		
		BCC644 = LMG2628	LT708134		
		BCC104	LT708132		
		BD316	LT708139		
		BD664	LT708133		
		BCC151	LT708135		
		SUPP1791	LT708144		
		y101	LT708131		
		BCC147 = CTB1135	LT708141		
		BCC601 = ICMP10132	LT708138		
		BCC155 = LMG2807	LT708140		
		<i>Pan. allii</i>	<i>atpD</i>	BCC632 = BD390 = LMG24248	LT708148
				BD391	LT708146
				BCC631 = BD377	LT708147
				BCC627 = BD309	LT708149
				BD381	LT708150
				BD304	LT708145
<i>Pan. ananatis</i>	<i>gyrB</i>			B1-9	NZ_CAEJ01000140
		BCC367	LT708151		
		BCC376	LT708152		
		LMG5342	PANA5342_RS22655		
		BCC132 = LMG2665 ^T	EF988824		
		BCC154 = LMG2664	LT708153		
		BCC149 = LMG2675	LT708154		
		BCC644 = LMG2628	LT708155		
		BCC104	LT708156		
		BD316	LT708157		
		BD664	LT708158		
		BCC151	LT708159		
		SUPP1791	LT708162		
		y101	LT708163		
		BCC147 = CTB1135	LT708164		
		BCC601 = ICMP10132	LT708160		
		BCC155 = LMG2807	LT708161		

Species	Locus	Isolate name	Accession Number/Locus Tag
<i>Pan. allii</i>		BCC632 = BD390 = LMG24248 BD391 BCC631 = BD377 BCC627 = BD309 BD381 BD304	LT708165 LT708166 LT708167 LT708168 LT708169 LT708170
<i>Pan. ananatis</i>	<i>infB</i>	B1-9 BCC367 BCC376 LMG5342 BCC132 = LMG2665 ^T BCC154 = LMG2664 BCC149 = LMG2675 BCC644 = LMG2628 BCC104 BD316 BD664 BCC151 SUPP1791 y101 BCC147 = CTB1135 BCC601 = ICMP10132 BCC155 = LMG2807	NZ_CAEJ01000140 LT708179 LT708172 PANA5342_RS41445 EF9889910 LT708171 LT708183 LT708180 LT708178 LT708184 LT708176 LT708175 LT708173 LT708174 LT708182 LT708177 LT708181
<i>Pan. allii</i>		BCC362 = BD390 = LMG24248 BD391 BCC631 = BD377 BCC627 = BD309 BD381 BD304	LT708187 LT708190 LT708188 LT708186 LT708189 LT708185
<i>Pan. ananatis</i>	<i>ompF</i>	B1-9 BCC367 BCC376 LMG5342 BCC132 = LMG2665 ^T BCC154 = LMG2664 BCC149 = LMG2675 BCC644 = LMG2628 BCC104 BD316 BD664 BCC151 SUPP1791	NZ_CAEJ01000140 LT708198 LT708192 PANA5342_2908 LT708205 LT708191 LT708195 LT708196 LT708203 LT708194 LT708201 LT708200 LT708199

