Taxonomy of *Pantoea* associated with bacterial blight of *Eucalyptus*

by

Carrie Louise Brady

Submitted in partial fulfilment of the
requirements for the degree
Magister Scientiae

in

The Faculty of Natural and Agricultural Sciences
University of Pretoria
Pretoria

Supervisor: Dr S.N. Venter
Co-supervisor: Prof T.A. Coutinho

Taxonomy of *Pantoea* associated with bacterial blight of *Eucalyptus*

by

Carrie Louise Brady

Supervisor: Dr S.N. Venter
Co-supervisor: Prof T.A. Coutinho
Department: Microbiology and Plant Pathology
Degree: MSc (Microbiology)

SUMMARY

The genus *Pantoea* has seven species and two sub-species, isolated from diverse geographical and ecological sources. The majority of *Pantoea* species are plant-associated and cause a wide variety of diseases on a range of hosts. *P. ananatis* causes disease on many agricultural crops including onion, maize, sudangrass, honeydew melon and pineapple. *P. ananatis* has been identified as the causal agent of a serious bacterial blight and dieback of *Eucalyptus* in South Africa. Bacterial isolates have also been recovered from Eucalypts in South America and Uganda exhibiting *Pantoea ananatis*-like symptoms. These isolates have not been identified. Identification of *P. ananatis*, based on phenotypic analysis, is difficult due to similarities in phenotypic characteristics between *Pantoea* species and related *Enterobacteriaceae*.

Regular isolations of *P. ananatis* have highlighted the need for a rapid, molecular-based identification technique for the pathogen. A PCR assay, based on amplification of a partial region of the 16S-23S ITS gene with species-specific primers, was evaluated for the rapid identification

of *P. ananatis*. Authentic strains of *P. ananatis* were included in the study, along with the unidentified isolates from South America and Uganda, and authentic strains of all species of the genus *Pantoea*. All authentic strains of *P. ananatis* produced a single PCR product of 398 bp following amplification. Only one unidentified isolate from South America produced the 398 bp PCR product. The only other species to be detected by the primers was *P. stewartii* subsp. *indologenes*. For the 16S-23S ITS-PCR assay to successfully detect only *P. ananatis*, the species-specific primers will have to be modified to increase their specificity.

Little is known concerning the genetic relatedness between species and strains of *Pantoea*, and no standardised molecular typing system exists for the genus. The entire 16S-23S ITS gene was evaluated for a genetic relatedness study for the genus *Pantoea*. Universal primers were used to amplify the entire spacer region. Multiple amplification products were visible for all *Pantoea* strains and the unidentified isolates. This indicated that the *Pantoea* genome contains multiple copies of the rRNA operon and a high degree of similarity exists among the rRNA operons of species of the genus *Pantoea*. Therefore it is not possible to determine the genetic relatedness of *Pantoea* species based on a typing technique targeting the 16S-23S ITS region.

The entire genome was then screened by AFLP analysis to examine the genetic relatedness of the genus *Pantoea*. The AFLP technique was found to be successful and distinct clusters were visible for each *Pantoea* species in the dendrogram. The majority of the South American and Ugandan isolates formed three separate clusters from *P. ananatis*. Representative strains were chosen from among the unidentified isolates for 16S rRNA sequencing. Based on the resulting phylogram, it is clear that at least two new *Pantoea* species or subspecies exist among the South American and Ugandan isolates, and more than one *Pantoea* species may be associated with bacterial blight and dieback of *Eucalyptus*. DNA-DNA hybridisations will be performed on these isolates to determine their correct taxonomic position within the genus *Pantoea*.

ACKNOWLEDGMENTS

- I would like to thank my parents, Garry and Yvonne, for giving me the opportunity to further my studies and for never holding me back. They have supported me unwaveringly in all endeavours and have always had confidence in me. For this I am extremely grateful, and blessed to have parents who care so much.

- To Alan, for his patience, support and love. Also for his assistance with all the computer graphics.

- All the students from the pathology lab, many of whom became friends. Lisa, Michelle and Sonya, thanks for all the laughs and good times.

- Dr S.N. Venter and Prof T.A. Coutinho, for your guidance, patience and advice. Thank you for the incredible opportunity which this project turned out to be.

- Prof M.J. Wingfield for collection of the South American isolates.

- Marc Vancanneyt and Ilse Cleenwerck for their kindness, advice and hospitality in Gent.

- DoL/NRF for the financial support over the past three years.

CONTENTS

LIST OF ABBREVIATIONS

AFLP	-	amplified fragment length polymorphism
ATP	-	adenosine triphosphate
bp	-	basepair
ºC	-	degrees Celsius
CHEF	-	contour-clamped homogenous electric field
CYVD	-	cucurbit yellow vine disease
dATP	-	deoxyadenine triphosphate
dCTP	-	deoxycytosine triphosphate
dGTP	-	deoxyguanine triphosphate
dTTP	-	deoxythymine triphosphate
DGGE	-	denaturing gradient gel electrophoresis
DNA	-	deoxyribonucleic acid
DTT	-	dithiothreitol
EDTA	-	ethylenediaminetetraacetic acid
EMBL	-	European Molecular Biology Laboratory

ERIC	-	enterobacterial repetitive intergenic consensus
fig.	-	figure
G + C mol %	-	moles percent guanosine plus cytosine
ITS	-	internally transcribed spacer
KAc	-	potassium acetate
kb	-	kilobase
KZN	-	KwaZulu/Natal
M	-	molar
Mb	-	megabase
MgAc	-	magnesium acetate
$MgCl_2$	-	magnesium chloride
mg/mL	-	milligrams/millilitre
MLEE	-	multilocus enzyme electrophoresis
µL	-	microlitre
µM	-	micromolar
MLST	-	multilocus sequence typing

mM	-	millimolar
ng	-	nanogram
PAGE	-	polyacrylamide gel electrophoresis
PFGE	-	pulsed-field gel electrophoresis
pmol	-	picomole
PCR	-	polymerase chain reaction
REP	-	repetitive extragenic palindromic
RFLP	-	restriction fragment length polymorphism
RNA	-	ribonucleic acid
rRNA	-	ribosomal ribonucleic acid
RS-HP	-	ribosomal spacer-heteroduplex polymorphism
R.S.A.	-	Republic of South Africa
SE-AFLP	-	single enzyme-AFLP
sp.	-	species
ssp.	-	subspecies
subsp.	-	subspecies

TAE	-	Tris.HCl-Sodium acetate-EDTA
TBE	-	Tris-Boric acid-EDTA
TEMED	-	N,N,N',N'-tetramethylethylene diamine
Tris HAc	-	Tris (hydroxymethyl)aminomethane acetate
U	-	unit
UPGMA	-	unweighted pair-group method with arithmetic mean
U.S.A.	-	United States of America
UV	-	ultraviolet
V	-	volt
W	-	watt

INTRODUCTION

Pantoea ananatis, formerly classified as *Erwinia ananas* (syn. *Erwinia uredovora*) (Mergaert *et al.*, 1993), belongs to the family *Enterobacteriaceae* and is a plant-associated pathogen. It is a Gram-negative, rod-shaped, yellow-pigmented bacterium. It is motile by means of a peritrichous flagellum and is non-spore-forming and facultatively anaerobic. Strains can grow at temperatures of 4-41 ºC, but no strain grows at 44 ºC. The type strain of *Pantoea ananatis* was isolated from pineapple in Brazil (Mergaert *et al.*, 1993). The name of this species has its origin in the generic name of pineapple, *Ananas comosus*, the source from which it was first isolated.

Pantoea ananatis has been identified as the causal agent of bacterial blight of *Eucalyptus* (Coutinho *et al.*, 2002); leaf spot of maize (Paccola-Meirelles *et al.*, 2001); necrotic leaf blotch disease of sudangrass (Azad *et al.*, 2000); leaf blight, seed stalk rot and bulb decay of onion (Gitaitis and Gay, 1997); postharvest disease of cantaloupe fruit (Bruton *et al.*, 1991) and brown spot of honeydew melons (Wells *et al.*, 1987). The pathogen has also been detected in naturally infested onion seed (Walcott *et al.*, 2002). The first reports of *P. ananatis* date back to 1928 where it was discovered to cause brown rot of pineapple fruitlets (Serrano, 1928).

In Georgia, U.S.A., *P. ananatis* has been recovered from 25 asymptomatic weed species, including crabgrass, Texas millet and tall verbena (Gitaitis *et al.*, 2002). Additionally the bacterium was also found on crop plants such as Bermuda grass, cowpea and soybean. The epiphytic presence of *P. ananatis* on weeds in Georgia implicates them as a possible source of inoculum. In the case of onion and sudangrass, it is known that the pathogen is both seedborne and seed-transmitted (Walcott *et al.*, 2002; Azad *et al.*, 2000)

The first report of *Pantoea ananatis* causing a disease on *Eucalyptus*, occurred in KwaZulu/Natal in 1998 (Coutinho *et al.*, 2002). Bacterial blight and dieback of the trees appeared in a single nursery on an *E. grandis* x *E. nitens* hybrid and spread to other nurseries and plantations, infecting a number of different *Eucalyptus* species and clones. Symptoms of bacterial blight include tip dieback and leaf spots on young leaves. *P. ananatis* appears to spread from the leaf petiole to the main vein and from there, infects the adjacent tissue. Initially, the leaf spots are

water-soaked and they often coalesce to form larger lesions. Leaf petioles become necrotic, resulting in premature abscission of the leaves. In the advanced stages of the disease, trees assume a scorched appearance and following repeated infection, become stunted (Coutinho *et al.*, 2002).

In South Africa, *Eucalyptus* species, hybrids and clones are grown commercially and account for more than 50 % of all newly afforested areas (Anonymous, 1996). Besides being used in the production of solid timber products, these trees form the basis of an internationally important pulp and paper industry. *P. ananatis* has the ability to infect a number of *Eucalyptus* species, hybrids and clones, which are crucial planting stock for the forestry industry, therefore this pathogen is of considerable concern. *Eucalyptus* trees suffering from bacterial blight in South Africa are more prevalent in areas where the temperature is low (20-25 °C) and the relative humidity is high. Strains have been recovered from Eucalypts showing similar symptoms in Uganda and three South American countries. The identity of these isolates has yet to be confirmed.

It is currently unknown as to how *P. ananatis* enters into its Eucalypt host. However, it has been established that the pathogen can enter other hosts through flowers, mechanical injury (Serrano, 1928), wounds created by feeding insects (Gitaitis *et al.*, 2003; Wells *et al.*, 1987) and plant-to-plant contact during high winds (Azad *et al.*, 2000). *P. ananatis* has been reported as a gut inhabitant of brown planthoppers (Watanabe *et al.*, 1996) and mulberry pyralids (Watanabe and Sato, 1999), however, neither of these insects act as a vector for the pathogen. Recently it was discovered that tobacco thrips do act as vectors for the transmission of *P. ananatis* to onions, causing the disease centre rot (Gitaitis *et al.*, 2003)

Serious economic losses can result from crops infected with *P. ananatis*. Blighted foliage can reach 50 % or more of the total leaf area of sudangrass infected with leaf blotch disease (Azad *et al.*, 2000). Leaf spot disease of maize results in a decrease in grain size and weight, and a decrease of 63,2 % in grain yield has been detected (Pinto, 1995). In the case of centre rot of onions, yield losses from 20-25 % to 80-100 % have been experienced (Gitaitis and Gay, 1997). *P. ananatis* is not only a primary plant pathogen in the field, but is also responsible for

postharvest losses of cantaloupe fruit (Bruton *et al.*, 1991), honeydew melons (Wells *et al.*, 1987) and onions (Gitaitis *et al.*, 2003).

Pantoea ananatis is one of only a few species of bacteria which includes ice-nucleating strains (Sigee, 1993). Bacteria exhibiting this trait can initiate the formation of ice that results in frost injury (Lindow, 1983). Frost sensitive plants, such as *Eucalyptus*, poplar and stone fruit tress, become more susceptible to freezing damage by the epiphytic presence of ice-nucleating *P. ananatis* strains. This ice-nucleation activity has been exploited by the food industry and by biological control specialists. Extracellular ice-nucleators from *P. ananatis* have been tested and applied in the freezing of foods in order to obtain the desired texture (Zasypkin and Lee, 1999) and in freeze-drying of foods (Watanabe and Arai, 1987). Ice-nucleating strains of *P. ananatis* have also been found to markedly reduce the cold hardiness of the mulberry pyralid larvae (Watanabe and Sato, 1999) and thus these stains have the potential as biological control agents of insect pests.

The epidemiology of plant diseases caused by *P. ananatis* is unclear. In South Africa, *P. ananatis* has only been isolated from *Eucalyptus*, onion and maize. It is not known if these outbreaks were caused by the same strain or how the bacterium could have spread from onions to *Eucalyptus* to maize. The recent appearances of bacterial blight and dieback on *Eucalyptus* in countries that have purchased seed from South Africa suggests seed transmission (unpublished data). There are also unconfirmed reports of the occurrence of *P. ananatis* on Eucalypts in South America and Uganda.

Identification of *P. ananatis*, based on phenotypic analysis, is difficult due to similarities in phenotypic characteristics between *Pantoea* species and related *Enterobacteriaceae* (Mergaert *et al.*, 1984; Verdonck *et al.*, 1987). A rapid, accurate molecular identification technique has yet to be developed for *P. ananatis*. It is important to understand the diversity and relationships, within and among related species as this leads to meaningful taxonomic classification, accurate identification, pathogen detection and epidemiology studies (Avrova *et al.*, 2002). Several molecular techniques, including DNA-DNA hybridisation (Mergaert *et al.*, 1993), 16S rRNA gene analysis (Coutinho *et al.*, 2002) and rep-PCR (unpublished results) have been performed on

isolates of *P. ananatis* and related species. However, a taxonomic study of the genus *Pantoea*, using a genomic fingerprinting technique such as AFLP or PFGE, has not yet been undertaken.

The genetic relatedness of *P. ananatis* strains must be examined and compared to other *Pantoea* species, in an attempt to trace the origin and spread of this pathogen. Whole genome analysis of *Pantoea* species may also allow the characterisation of the unidentified South American and Ugandan isolates.

Aim:

To examine the taxonomy and relatedness of *Pantoea* strains recovered from Eucalypts in South Africa, South America and Uganda

Objectives:

- To develop a rapid, reliable molecular-based identification technique for *Pantoea ananatis*
- To evaluate a genomic fingerprinting technique for the genus *Pantoea* to examine the genetic relatedness between species and strains

MOLECULAR TYPING METHODS FOR USE IN THE CLASSIFICATION OF *PANTOEA* SPECIES

2.1) Introduction:

The family *Enterobacteriaceae* was formed to include all facultatively anaerobic, Gram-negative rods. The organisms described in this family are typically straight, rod-shaped bacteria and most are motile by means of a peritrichous flagellum. They grow well at 37 ºC and are oxidase negative and catalase positive, with few exceptions (Holt *et al.*, 1994). The majority of organisms placed in the *Enterobacteriaceae* are associated with the digestive tract and human disease, for example *Escherichia*, *Salmonella* and *Shigella*. However, phytopathogenic bacteria, including *Erwinia*, *Brenneria* and *Pectobacterium* are also found in this family. A recent addition to the family *Enterobacteriaceae* is the genus *Pantoea*, proposed by Gavini *et al.* in 1989.

In 1920, Winslow and co-workers first described the genus *Erwinia* (named after the phytobacteriologist Erwin F. Smith), under which all the Gram-negative, fermentative, non-sporulating peritrichous flagellated plant pathogenic bacteria were placed. In 1972, it was proposed that the "herbicola group" of *Erwinia* be placed under the name *Enterobacter agglomerans.* This included several *Erwinia* species, including *E. herbicola, E. milletiae, E. ananas, E. uredovora* and *E. stewartii* (Ewing and Fife, 1972).

Synonymy of the species names *Enterobacter agglomerans*, *Erwinia herbicola* and *Erwinia milletiae* was suspected, based on a high degree of similarity of protein electropherograms and phenotypic data (Verdonck *et al.*, 1987). The type strains of the *Erwinia herbicola-Enterobacter agglomerans* complex belonged to the same DNA hybridisation group, which confirmed the suspected synonymy (Beji *et al.*, 1988). It was proposed that a new genus, *Pantoea*, should be established consisting of two species: *Pantoea agglomerans* as the type species, including the type strains of

Enterobacter agglomerans, *Erwinia herbicola* and *Erwinia milletiae*, and a new species named *Pantoea dispersa* (Gavini *et al.*, 1989).

In Japan in 1992, bacterial strains isolated from fruit and soils samples shared the general characteristics of the family *Enterobacteriaceae*. Following DNA hybridisation, three new species were described and classified in the genus *Pantoea* as *P. citrea, P. punctata and P. terrea* (Kageyama *et al.*, 1992). A year later it was proposed to transfer *Erwinia ananas*, *Erwinia uredovora* and *Erwinia stewartii* to the genus *Pantoea* following DNA relatedness studies and fatty acid analysis. *E. ananas* and *E. uredovora* were united as a single species which was classified as *Pantoea ananatis*. Two separate subspecies were created within the species *Pantoea stewartii* (formerly *Erwinia stewartii*), *P. stewartii* subsp. *stewartii* and *P. stewartii* subsp. *indologenes* (Mergaert *et al.*, 1993). Currently, seven species and two subspecies have been described under the genus *Pantoea*.

Pantoea stewartii subsp. *stewartii* (Mergaert *et al.*, 1993) is the causal agent of Stewart's vascular wilt of sweet corn and maize, which can result in huge crop losses. The potential risk of seed transmission of *P. stewartii* subsp. *stewartii* is considered so important in the international shipment of corn seed, that over 50 countries prohibit its import unless it is certified to be free of the pathogen (Coplin *et al.*, 2002). In many countries *P. stewartii* subsp. *stewartii* is classified as a quarantine pathogen.

Pantoea ananatis is the next most devastating species in the genus, causing a number of diseases on a wide variety of plant hosts. This pathogen causes bacterial blight and dieback of Eucalypts (Coutinho *et al.*, 2002); leaf spot of maize (Paccola-Meirelles *et al.*, 2001); necrotic leaf blotch disease of sudangrass (Azad *et al.*, 2000); leaf blight, seed stalk rot and bulb decay of onion (Gitaitis and Gay, 1997); postharvest disease of cantaloupe fruit (Bruton *et al.*, 1991) and brown spot of honeydew melons (Wells *et al.*, 1987). The first reports of *P. ananatis* date back to 1928 where it was discovered to cause brown rot of pineapple fruitlets (Serrano, 1928). Up to 100 % crop losses have been experienced for the majority of the cases listed above. There is also a recent report of *P. ananatis* causing a bacteremic infection in an elderly patient (DeBaere *et al.*, 2004).

Pantoea agglomerans causes palea browning of rice (Xie, 2001), crown and root gall of the ornamental plant, *Gypsophila paniculata* (Cooksey, 1986) and is associated with millet, apples and beans. *P. citrea* causes pink disease of pineapple and has found to be associated with citrus fruit (Pujol and Kado, 2000). Although *P. terrea* and *P. punctata* have been isolated from citrus fruit, they have yet to be reported to cause disease on these plant hosts (Kageyama *et al.*, 1992).

Little is known concerning the specificity and genetic relatedness of *P. ananatis* strains from different hosts and geographical regions. Additionally the epidemiology of the diseases caused by this pathogen is unclear. It is known that the pathogen is seedborne and seed-transmitted in several hosts, and is also present as an epiphyte on some crops and weeds. It is possible that *P. ananatis* exists as an epiphyte on many plant hosts and only causes disease under certain conditions.

As *Pantoea* is a recently described genus, there is still much confusion regarding taxonomic nomenclature. *Erwinia ananas* and *Erwinia uredovora* were synonymized based on a high level of genetic relatedness in 1993 (Mergaert *et al.*, 1993). However, the synonymy of these two species is not widely accepted, and the name *E. uredovora* is still commonly used in literature (Beyer *et al*., 2002). Similarly, the species names *Enterobacter agglomerans* and *Erwinia herbicola* are still used (Muniruzzaman *et al.*, 1994; Nizan-Koren *et al.*, 2003), despite the inclusion of these names under *Pantoea agglomerans* (Gavini *et al.*, 1989). In the most recent edition of Bergey's Manual of Determinative Bacteriology, only two species of *Pantoea* are listed, *Pantoea agglomerans* and *Pantoea dispersa* (Holt *et al.*, 1994). There is a need for the reference material to be updated to set a standard for the use of the correct nomenclature for *Pantoea* and its species.

Pantoea may only be a small genus, with seven known species, but its range of hosts is extremely varied. Species of *Pantoea*, especially *Pantoea ananatis*, are isolated from, or found to cause disease on, new hosts almost every year. Despite these regular isolations, there is little sequence data available for specifically *P. ananatis, P. dispersa, P. citrea, P. punctata* and *P. terrea*. This needs to be remedied if the identification and characterisation of these pathogens is to be rapid and accurate. The purpose of this review is to consider the various molecular typing methods previously applied to bacteria of the family *Enterobacteriaceae*; and to determine which of these methods can be utilised for a taxonomic study of the genus *Pantoea*.

2.2) Microbial Typing:

The term "typing" is used to indicate the differentiation of strains at the subspecies level or below and can be either identifying or comparative. Identifying, or determinative typing is utilised in the allocation of organisms to a previously described type in an existing classification scheme. With comparative typing, microbial isolates of a defined set are compared to each other for similarity, with reference to existing classification schemes (Dijkshoorn and Towner, 2001).

A high degree of similarity exists, not only among the species of the phytopathogenic enterobacteria, but between all organisms within the family *Enterobacteriaceae*. This is illustrated by close resemblance of their phenotypic characteristics, chemical composition and conservative 16S rRNA sequences (Stackebrandt *et al.,*1999). In the past phenotypic methods such as fatty acid composition and substrate utilisation profiles were used for the identification and classification of bacteria. However, phenotypic methods have low reproducibility and may not reflect the genetic relatedness of bacterial isolates. Additionally, phenotypic characteristics of microorganisms are susceptible to change depending on growth conditions, growth phase and spontaneous mutation (Wilson and Spencer, 1999). In contrast, DNA-based methods, specifically those that are PCR-based, are more stable, can define genetic relatedness clearly and are therefore better suited for microbial typing (Louws *et al.*, 1999).

2.2.1) Detection and Identification:

Detection involves establishing the presence of a particular target organism within a sample, whilst diagnosis refers to the identification of the nature and cause of a disease problem (Shurtleff and Averre, 1997). Techniques for both detection and diagnosis require certain levels of specificity, sensitivity and speed (Louws *et al.*, 1999). In the case of PCR-based methods, specificity is determined by primer selection and amplification conditions, while sensitivity depends on the nature of the PCR protocol and sample. Identification is defined as the assignment of an unknown organism into a known taxonomic group, based on selected characteristics (Shurtleff and Averre, 1997). PCR-based identification techniques, for detection and diagnosis of pathogens, involve the use of specific primers to amplify a diagnostic fragment or universal primers that provide a diagnostic genomic fingerprint (Louws *et al.*, 1999).

2.2.2) Diversity and Phylogeny:

Diversity refers to the degree of genetic variation within bacterial populations and is related to bacterial systematics at multiple taxonomic and phylogenetic levels (Louws *et al.*, 1999). Taxonomy can be divided into three parts: classification, nomenclature and identification of unknown organisms (Vandamme *et al.*, 1996). Diversity exists even at the species level, meaning that organisms isolated from different sources at different times and in different geographical regions can be classified into subtypes or strains (Olive and Bean, 1999). It is necessary to determine the genetic diversity of bacterial populations to establish a stable taxonomy, which leads to better classification schemes. Phylogenetics is the process of reconstructing possible evolutionary relationships and uses nucleotide sequences from conserved genes that act as molecular markers (Owen, 2004). The use of small-subunit RNA genes in the construction of branching trees, representing the distance of divergence from a common ancestor, has made a vast impact on microbial phylogenetics.

However, it is DNA-DNA hybridisation that is considered the "gold standard" for the delineation of bacterial species (Stackebrandt and Goebel, 1994). A species is defined as a group of strains, including the type strain, sharing 70 % or greater DNA-DNA relatedness with 5 ºC or less melting temperature of the hybrid (Wayne *et al.*, 1987). The percent DNA binding is an indirect parameter of the sequence similarity between two entire genomes. DNA hybridisation techniques have severe disadvantages as they depend on physiochemical parameters, are not cumulative, are cumbersome and require large quantities of DNA (Stackebrand and Goebel, 1994). Despite these disadvantages, DNA-DNA hybridisation values are considered part of the standard description of bacterial taxa, along with the moles percent guanosine plus cytosine (Vandamme *et al.*, 1996).

2.2.3) Molecular Typing Techniques:

There are several criteria which molecular typing methods must meet in order to be broadly useful. Firstly, a molecular typing method should require no previous knowledge of DNA sequence. Secondly, all organisms within a species must be typeable by the method used. Thirdly, any typing method must have a high discrimination power and lastly, a typing method has to provide reproducible results which can be easily interpreted (Olive and Bean, 1999). The choice of typing method depends on technical difficulty, cost, laboratory resources and length of time needed to obtain a result.

Molecular typing techniques can be characterised into two categories: specific gene variation and random whole-genome analysis. Specific gene variation includes single-locus and multi-locus typing. Highly variable genes, such as pathogenicity islands, can be targeted by the single-locus approach as they have been directly implicated in causing disease (Gürtler and Mayall, 2001). Pathogenicity islands have been found in a number of bacterial species, and contain virulence genes and mobilisation elements. However, the high variability of single loci chosen for molecular typing often obscures true relationships between isolates (Gürtler and Mayall, 2001). Multilocus sequence typing (MLST) and analysis of multi-gene families, such as the RNA operons and tRNA genes, are examples of the multi-locus approach to molecular typing. MLST involves

sequencing of several housekeeping alleles where the unique sequence of each gene fragment is considered a unique allele (Gürtler and Mayall, 2001). The rRNA operon contains three rRNA genes (16S, 23S and 5S) and three spacer regions which may include a tRNA gene. 16S rRNA sequencing is considered commonplace for phylogenetic analyses, whilst the 23S rRNA gene and tRNA genes have only been used to a limited extent. However, it is the 16S-23S internally transcribed spacer (ITS) region which is being used more frequently for molecular typing and microbial diversity studies.

Random whole-genome analysis techniques recognise random sites on the genome, that cannot be predicted without the whole genome sequence (Gürtler and Mayall, 2001). These random sites include repetitive elements and restriction enzyme sites. Repetitive extragenic palindromic-PCR (Rep-PCR) targets repetitive elements which are scattered throughout the genome, whilst pulsed-field gel electrophoresis (PFGE) and amplified fragment length polymorphism (AFLP) analysis are used to separate DNA fragments produced by digestion with restriction enzymes. Rep-PCR, PFGE and AFLP all generate DNA banding patterns, or genomic fingerprints, based on the genomic sequence of a microorganism. A genomic fingerprint can be described as a display of a set of DNA fragments from a specific DNA sample. When the entire genome is screened by random whole-genome analysis, variations in banding patterns are a direct reflection of the genetic relationship between the bacterial strains examined. Genomic fingerprints allow numerical analysis for comparative typing and identification purposes (Janssen *et al.*, 1996). In bacteriology, discrimination to the species level is most often referred to as identification, whilst typing denotes differentiation to the strain level (Savelkoul *et al.*, 1999). AFLP analysis can be used for both identification and typing of microorganisms.

From the two molecular typing categories, amplification of the 16S-23S ITS region and AFLP- or PFGE-analysis meet the most requirements stated previously. Neither require previous DNA sequence knowledge; both are reproducible and yield results that can be easily interpreted. These techniques will be examined in more depth in this review

2.3) Specific Gene Variation:

2.3.1) The 16S-23S Internally Transcribed Spacer (ITS) Region:

The rRNA operon is of great evolutionary interest since it is found in both eukaryotic and prokaryotic organisms. The rRNA operon, unlike metabolic operons (eg. *lac* operon and *trp* operon), does not encode for any proteins (Adhya, 1999). In prokaryotes, the rRNA operon consists of three conserved genes: the small subunit 16S rRNA gene, the large subunit 23S rRNA genes and the 5S rRNA gene. These functionally conserved genes are separated by variable spacer regions. The prokaryotic 16S rRNA gene has been used extensively in taxonomy studies at and above the species level. Because this gene has highly conserved regions interspersed with variable and hypervariable sequences, primers can be easily designed to amplify these sequences. Additionally a large database of 16S rRNA sequences is available for comparisons.

One of the disadvantages of targeting the 16S rRNA gene in diversity studies is that the length of the gene is considerably constant. Therefore amplified 16S rRNA genes from different isolates cannot be separated easily by electrophoresis. And, despite containing variable and hypervariable regions, they are often not divergent enough to distinguish between strains within the same species and sometimes even between species of the same genus (García-Martínez *et al.*, 1999). The length of the 23S rRNA gene sequence is twice as long as that of the 16S gene and relatively few sequences have been submitted to the ribosomal database. Consequently neither the 23S gene nor the 5S gene have been widely used in phylogenetic studies (Gürtler and Stanisich, 1996).

The spacer regions that separate the rRNA genes, do however exhibit a large degree of sequence and length variation at the genus and species level. The size of the spacer may vary among different species and even among different operons within a single cell, in the case of multiple operons (Condon *et al.*, 1995). In bacterial genomes, up to 11 copies of the rRNA operon can be present within the chromosome. After transcription of the rRNA operon, one pre-rRNA transcript is present which contains the following components:

5' – 16S gene – spacer – tRNA – spacer – 23S gene – spacer – 5S gene – 3'

This pre-rRNA transcript is then cleaved into separate rRNA and tRNA molecules (Watson *et al.*, 1987). The spacer regions found between the rRNA genes within the operon are more diverse due to the variation in number and type of tRNA sequences located within these spacers. The majority of spacer regions in Gram negative bacteria contain both a $tRNA^{ala}$ and a $tRNA^{ile}$ gene, whereas the spacers in Gram positive bacteria rarely contain any tRNA genes (Gürtler and Stanisich, 1996). Other functional units within the spacer region include the recognition sequence for the enzyme ribonuclease III, which is involved in splicing to produce the mature ribosome (Bram *et al.*, 1980).

The internally transcribed spacer (ITS) region, situated between the 16S and 23S ribosomal genes has been used extensively for both identifying and comparative typing of microbes. This region consists of a series of conserved, alignable sequences, found in all strains of a single species, alternating with stretches of hypervariable sequences that are of equal length or that contain insertions or deletions (García-Martínez *et al.*, 2001). Precise species identification can be achieved using the alignable sequences of the ITS region, whilst the hypervariable sequences often allow for characterization below the species level.

Techniques:

The ITS region is ideally situated for easy PCR amplification and sequencing with conserved regions of the adjacent 16S and 23S genes acting as primer-binding sites. There are several techniques that can be applied to the PCR amplification product to reveal the polymorphisms present in the ITS region.

ITS-PCR:

When universal primers are utilised, identification of isolates can be based on the number and length of ITS regions amplified (Jensen *et al.*, 1993). The occurrence of more than one band generated from a single isolate can result from varying lengths of several copies

of the rRNA operon. The variation among operons within a multi-rRNA operon genome may be as great as the variation of operons between closely related strains (Nagpal *et al.*, 1998).

Primer design, for the amplification of the ITS gene, is made simple when the sequences of the adjacent 16S and 23S rRNA genes are known. Also, PCR conditions can be optimised to increase efficiency of amplification. ITS-PCR has been successfully used in the detection of *P. ananatis* from weeds and crops (Gitaitis *et al.*, 2002) and for a phylogenetic study of phytopathogenic *Enterobacteriaceae* (Hauben *et al.*, 1998).

PCR-RFLP:

PCR-restriction fragment length polymorphism (PCR-RFLP) analysis combines specific PCR-based amplification, usually of a gene or part of a gene, with restriction digestion of that PCR product which is separated electrophoretically. If the restriction enzyme recognition site is uniquely situated, the resultant fragment size pattern can be used to differentiate between species and strains of organisms (Jensen *et al.*, 1993). PCR-RFLP analysis of the 16S-23S spacer region was used in subspecies differentiation of *Salmonella choleraesuis* (Shah and Romick, 1997). However, this technique relies on the presence or absence of a restriction enzyme site and as most mutations do not result in the creation or destruction of restriction enzyme sites, PCR-RFLP is not always a viable option for typing of isolates.

Applications:

Detection and Identification:

In order to facilitate genus and species level identification of bacteria, Jensen and co-workers developed a set of primers and PCR conditions to amplify the 16S-23S ITS region. Strains of *Salmonella, Citrobacter, Yersinia, Enterobacter, Escherichia* and *Proteus* from the family *Enterobacteriaceae* were screened using universal primers (Jensen *et al.*, 1993). The banding patterns produced were distinct for each different

species of the genera. Following comparison of the pattern profiles with a database, it was found that common profiles exist for each species, and it is possible to distinguish between the bacterial isolates screened, based on these profiles.

Seven different rice seed pathogens of the genera *Erwinia*, *Pseudomonas* and *Xanthomonas* were detected based solely on size polymorphisms of the primary and secondary ITS-PCR products generated by the use of universal primers complementary to the 16S and 23S rRNA genes (Kim and Song, 1996). A new species, *Erwinia pyrifoliae*, was proposed after isolates from Asian pears were analysed by sequencing the 16S rRNA gene and the ITS region (Kim *et al.*, 1999). The 16S rRNA gene is almost identical to that of *Erwinia amylovora*, however the ITS region sequence revealed that the Asian pear pathogen is only closely related to *E. amylovora*. DNA hybridisation results and the G + C content were significantly different to *E. amylovora* and a new species was suggested.

An alternative to universal primers, is the design of species-specific primers which were used in the detection of *Pantoea ananatis* as the causal agent of centre rot of onion (Walcott *et al.*, 2002). The 16S-23S ITS regions, of isolates from infected onions, were amplified using species-specific primers, designed based on nucleotide sequences of the ITS region of a known *P. ananatis* isolate. The presence of a single 400 bp PCR product is a positive indication of *P. ananatis* or *P. stewartii*, as the primers amplify the ITS regions of both of these species.

Bacterial isolates from blackleg-infected potatoes in Brazil were identified as pectolytic *Erwinia* species following amplification and sequencing of the 16S-23S ITS region (Duarte *et al.*, 2004). RFLP analysis of the ITS region differentiated the Brazilian strains from the subspecies of *E. carotvora* and from *E. chrysanthemi*. This confirmed that blackleg disease of potato is caused by a different strain of pectolytic *Erwinia* in Brazil than in temperate potato-growing regions.

A PCR assay with universal primers, specific for the 16S-23S ITS region, was evaluated for the detection and identification of waterborne pathogens such as *Escherichia coli*, *Salmonella enterica*, *Shigella dysenteriae* and *Yersinia enterolitica* (Pérez-Luz *et al.*, 2004). Specific profiles for each of the species were obtained, but the technique could not simultaneously detect microbes present in mixed population samples. However, a possible rapid identification method was found for waterborne bacteria.

Diversity and Phylogeny:

Up to eleven copies of the rRNA operon can be present on a bacterial genome. *Escherichia coli* K-12 has seven copies, located near to the origin of replication. PCR-RFLP analysis of the 16S-23S ITS region of *E. coli* strains revealed two different clusters present in the species, α and β (García-Martínez *et al.*, 1996[a]). To detect the sequence divergence between the two clusters, the 16S-23S ITS regions were sequenced and four phylogenetically informative variable sites were identified. Comparison of the sequences confirmed that two major phylogenetic branches exist within the species *E. coli* (García-Martínez *et al.*, 1996[b]). Comparison of sequences from different copies of rRNA operons from the same strain, revealed a degree of intercistron heterogeneity, meaning that identical rRNA operons are rarely found on the same genome.

To investigate the diversity of *Salmonella* strains isolated from different natural aquatic systems, strains were screened using the ribosomal spacer-heteroduplex polymorphism (RS-HP) method (Baudart *et al.*, 2000). Because seven copies of the rRNA operon exist in the *Salmonella* genome, amplification of the 16S–23S ITS region results in multiple amplification products. When multiple amplification products, containing homologous sequences flanking heterologous intervening sequences are generated, the products can cross-hybridize to form heteroduplex DNA structures (Jensen and Hubner, 1996). It was found that all serotypes of *Salmonella* can be easily differentiated using the RS-HP profiles and that a large diversity of serotypes exist within the natural aquatic environment.

Denaturing gradient gel electrophoresis (DGGE) was applied to the 16S-23S ITS region as a means to evaluate strain level differences in *E. coli* (Buchan *et al.*, 2001). DGGE relies on polyacrylamide gels that have an increasing linear gradient of denaturants (formamide and urea) to separate double-stranded DNA fragments on the basis of their sequence, not size (Fischer and Lerman, 1979). The sensitivity of DGGE to any slight sequence differences is high, meaning that single base changes are often revealed. The application of DGGE to the internally transcribed spacer region proved to be an effective method for diversity studies on field-collected samples, as unique banding patterns were obtained for *E. coli* isolates from each different source.

Sequence analysis of the 16S-23S ITS region was evaluated for phylogenetic studies of *E. amylovora* and *E. pyrifoliae* (McGhee *et al.*, 2002). It was found that the use of the ITS region for phylogenetic studies is questionable due to the high degree of variability within the spacer regions. Sequence analysis of the spacer regions of *E. amylovora* and *E. pyrifoliae* implied that recombination occurred between copies of the spacers, from different operons on the same chromosome. Therefore, misleading results can be produced when applying phylogenetic methods, meant for single nucleotide mutational substitutions, to polymorphisms due to recombination events (McGuire *et al.*, 1997).

2.4) Random Whole-Genome Analysis:

2.4.1) Repetitive Extragenic Palindromic-PCR (Rep-PCR):

Technique:

Rep-PCR is based on amplification from the sites of repetitive extragenic palindromic (REP) elements, enterobacterial repetitive intergenic consensus (ERIC) sequences and BOX elements found at different positions on the bacterial genome. The term "rep-PCR" includes amplification of any of the repetitive elements. REP elements are 38 bp long sequences consisting of six degenerate positions and a five bp variable loop between each side of a conserved palindromic stem (Stern *et al.*, 1984).

If two REP elements are located close enough, they can serve as primer binding sites to amplify the region of DNA between them. These introns are called inter-repeat fragments and, depending on the primers used, the number of fragments amplified will vary between strains (Towner and Grundmann, 2001).

ERIC sequences are 126 bp long elements which contain a highly conserved central inverted repeat and are located in extragenic regions of genomes of the *Enterobacteriaceae* (Hulton *et al.*, 1991). The third set of repetitive elements are the BOX elements. These sequences are less commonly used for microbial typing than REP and ERIC sequences. BOX elements are located within intergenic regions and are mosaic repetitive elements composed of various combinations of three subunit sequences referred to as boxA, boxB and boxC (Martin *et al.*, 1992).

Amplification of rep elements can be performed with a single primer, a single set of primers or multiple sets of primers. High annealing temperatures are used in the PCR cycle to produce distinct DNA fingerprints at a relatively high stringency (Towner and Grundmann, 2001). The fingerprints generated by REP-PCR are more complex than those by ERIC-PCR, however both are discriminative at the strain level.

Rep-PCR is one of the most widely used methods for microbial typing. The technique is simple, rapid and can be applied to any number of isolates. Rep-PCR has considerably better discriminatory power than RFLP analysis of the 16S rRNA gene or the 16S-23S ITS region. However, primer annealing to other homologous DNA, besides rep elements cannot be discounted (Gillings and Holley, 1997) and rep-PCR may not necessarily differentiate between very closely related strains (Louws *et al.*, 1999).

Applications:

Rep-PCR was used to establish the genomic diversity of *E. coli* water isolates collected from southern Brazil (dos Anjos Borges *et al.*, 2003). Following analysis, 28 clusters and 16 single patterns were contained in the dendrogram. Genetic diversity between isolates from different sampling sites could be observed from the results. *Erwinia* soft-rot bacterial isolates, from ornamental plants in Florida were characterized by Rep-PCR (Norman *et al.*, 2003). Strains could be clustered into two heterogeneous populations of *E. chrysanthemi* and *E. carotovora* subsp. *carotovora*, which correlated with fatty acid analysis and biochemical test results.

The identity of the bacterium that causes cucurbit yellow vine disease (CYVD) was confirmed as *Serratia marcescens* by Rep-PCR (Zhang *et al.*, 2003). CYVD strains were compared to strains of *S. marcescens* from different niches and type strains of other *Serratia* species. CYVD strains formed a loosely related group with *S. marcescens* strains from other niches, confirming its identity. Pulsed-field gel electrophoresis (PFGE) and Rep-PCR, using BOX, ERIC and REP primers, were compared for the genomic detection for *Salmonella* isolates (Weigel *et al.*, 2004). Both techniques were highly reproducible and produced approximately the same number of DNA fragments. However the Rep-PCR profiles were more variable and allowed greater discrimination.

2.4.2) Pulsed-Field Gel Electrophoresis (PFGE):

Technique:

Macrorestriction analysis of microbial genomes resolved by pulsed-field gel electrophoresis (PFGE) emerged in the 1990's as the method of choice for epidemiological studies, and is considered the "gold-standard" of molecular typing methods (Struelens *et al.*, 2001). Macrorestriction fingerprinting refers to digestion of chromosomal DNA with infrequent cutting enzymes, whilst PFGE implies gel electrophoresis where the electric field periodically changes in direction and intensity.

This is a high resolution technique which allows separation of DNA fragments as large as 12 Mb.

Intact, whole chromosomal DNA is required for macrorestriction analysis. This is achieved by incorporating bacterial isolates into ultra-pure, nuclease-free, low-melting temperature agarose plugs. The agarose plugs, containing whole bacteria, are then exposed to lytic solutions and infrequently-cutting restriction enzymes. Infrequently-cutting enzymes are selected on the basis of the rarity of occurrence of their recognition sequence in the target genome. Endonucleases which include CTAG in their recognition sequence cleave high G + C content genomes less than once per 1 kb, for example *Xba*I (TCTAGA) and *Spe*I (ACTAGT).

The digested bacterial plugs are inserted into an agarose gel and subjected to pulsed-field gel electrophoresis. The contour-clamped homogenous electric field (CHEF) system is the most popular PFGE system used for bacterial typing. A hexagonal array of electrotrodes periodically alternates uniform fields with an angle of reorientation of 120 ° (Struelens *et al.*, 2001). In response to changes in the orientation of the electric field, large DNA molecules migrate through the gel in a zig-zagging motion.

DNA fragments are visualized after ethidium bromide staining of the gel, followed by digital image capture under UV light. Computer-assisted analysis of PFGE gels allows rapid identification of strains by comparison of patterns to those in a large database. Comparison of patterns can be based either on densiometric curve comparisons of the Pearson correlation coefficient, or on the fragment matching Dice coefficient of similarity. The Dice coefficient, calculated as the number of matching size fragments multiplied by two and divided by the total number of fragments in a pair of patterns, is the preferred coefficient for comparing PFGE patterns (Struelens *et al.*, 2001).

A system for standardising the interpretation of PFGE patterns has been proposed, in which bacterial isolates yielding the same PFGE pattern are considered the same strain. Isolates differing by one to three bands are closely related and isolates differing by four to

six bands are possibly related. Bacterial isolates containing six or more band differences are considered unrelated (Tenover *et al.*, 1995).

Although macrorestriction analysis by PFGE has been used extensively for comparative typing in epidemiological investigations, it is still a laborious technique. The one major drawback of PFGE typing is the time needed to complete the analysis, which can take up to three days (Olive and Bean, 1999). Further automation of the technique would enhance the practicability, throughput and speed of the analysis (Struelens *et al.*, 2001).

Applications:

PFGE was performed on *Pantoea stewartii* subsp. *stewartii* isolates and patterns were compared to those of other *Pantoea* and *Erwinia* species (Coplin *et al.*, 2002). *P. stewartii* isolates could be easily distinguished based on their PFGE fingerprints. Sufficient divergence existed in the profiles to differentiate between *P. stewartii* subsp. *stewartii* isolates from different geographical regions. *Yersinia pestis* strains and isolates were screened by PFGE in an epidemiological study and all produced similar banding patterns (Huang *et al.*, 2002). Minor differences in the PFGE patterns occurred if the isolates came from different parent strains. PFGE offers an increased ability to discriminate between *Y. pestis* strains, thereby improving epidemiological studies.

Two different species of *Shigella*, *S. sonnei* and *S. flexneri*, were analyzed by PFGE typing. The genetic diversity of *Shigella sonnei* was examined by PFGE, ribotyping and plasmid profiling (DeLappe *et al.*, 2003). Two major clusters were identified by PFGE and a level of correlation existed between the different molecular typing methods. To trace the source of infection of a *Shigella flexneri* outbreak, PFGE was applied to all isolates recovered (Chiou *et al.*, 2001). Analysis of the PFGE patterns proved that *Shigella* infections were endemic to that area. PFGE analysis was used in epidemiological studies of outbreaks of *Klebsiella pneumoniae* (Ben-Hamouda *et al.*, 2002) and *Enterobacter cloacae* (Talon *et al.*, 2004) in two different neonatal wards.

However, in the case of the *E. cloacae* outbreak, the cause of the emergence of the pathogen could not be identified.

2.4.3) Amplified Fragment Length Polymorphism (AFLP) Analysis:

Technique:

The AFLP technique (Vos *et al.*, 1995), named for it's resemblance to restriction fragment length polymorphism (RFLP), combines the power of RFLP with the flexibility of PCR. AFLP is based on the detection of genomic restriction fragments by PCR amplification and can be applied to DNA from any origin. The technique consists of three steps: digestion of genomic DNA with two restriction endonucleases and ligation of double-stranded adaptors to the resulting restriction fragments; pre-amplification and selective amplification of restriction fragments with two sets of primers, complimentary to the ligated adaptors, and lastly, electrophoretic separation of the amplified products.

The two restriction endonucleases utilised in the AFLP technique include an infrequent (6 bp) cutter, and a frequent (4 bp) cutter. This ensures that a limited number of fragments are obtained which are of a relatively small size, simplifying the amplification process. For a microbial genome with a low G + C content (< 40 mol %), the restriction enzymes *Eco*RI and *Mse*I are recommended whilst *Apa*I and *Taq*I are suited for genomes with a high G + C content (> 55 mol %). For genomes with an intermediate G + C content (40 – 50 mol %), *Hin*dIII and *Taq*I could be used for restriction enzyme digestion (Janssen *et al.*, 1996).

Double-stranded adaptors, with sticky ends complimentary to those generated by the restriction enzyme digestion, are ligated to the restriction fragments in the presence of T4 DNA ligase and ATP. The adaptors are designed in such a way that the original restriction site is not restored after their ligation to the restriction fragments (Zabeau and Vos, 1993). This allows restriction enzymes to be present during ligation to prevent fragment-to-fragment ligation. As adaptors are unphosphorylated at their 5' ends,

adaptor-to-adaptor ligation is also prevented. The adjacent adaptor and restriction site sequences serve as primer binding sites for the subsequent amplification reactions.

AFLP primers consist of a 5' sequence corresponding to the adaptor, the restriction site sequence and a 3' selective extension. Therefore the design of the primers depends on the adaptors used during ligation. A pre-amplification reaction of the ligation mixture is performed, using primers with no selective nucleotides attached at their 3' ends. This pre-amplification step reduces the complexity of the ligation mixture and prevents background smears from appearing on the gel. Only restriction fragments with an adaptor ligated to each end will be amplified exponentially during pre-amplification.

The pre-amplification reaction is diluted and serves as a template for selective amplification. One to three selective nucleotides are included at the 3'-ends of the PCR primers, amplifying only a subset of the pre-amplified restriction fragments. Only restriction fragments, in which the nucleotides flanking the restriction sites are complementary to the selective nucleotides on the primers, will be amplified. As the number of selective nucleotides added to the primers increases, the complexity of the DNA fingerprint decreases. The number of amplified fragments is reduced four-fold with each additional selective nucleotide (Vos *et al*., 1995). At least three selective nucleotides are required at the 3'-ends of both primers to obtain suitable AFLP fingerprints for highly complex genomes, like those of plants or animals. For smaller genomes, like those of bacteria, usually only one selective nucleotide at the 3'-end of each primer is necessary (Vos *et al.*, 1995).

Mismatching between the primers and the template is not expected as AFLP primers are complementary to their target site of adaptor and restriction site sequences. The only possible mismatch that may occur is between the selective nucleotides at the 3'-end of the primer and the corresponding nucleotides flanking the restriction site (Kwok *et al.*, 1990). However, due to high-stringency PCR conditions, fragments with mispaired 3'-ends primers will not be amplified. The AFLP selective amplification begins with a high annealing temperature, which is lowered by 1 ºC per cycle until the optimal annealing

temperature is reached. This step is referred to as a "touchdown" and ensures the dissociation of mismatched primers and fragments, improving the specificity of the amplification.

Following selective amplification, reaction products are mixed with formamide loading dye and separated electrophoretically. In the past, selective primers were labelled radioactively and the resulting amplification products analysed on denaturing polyacrylamide gels. Autoradiography or phosphoimage analysis was used to visualise banding patterns (Vos *et al.*, 1995). However, as many laboratories prefer to limit the use of radioactivity, alternative detection methods have been evaluated, including silver-staining (Chalhoub *et al.*, 1997). Although silver-staining produces a similar resolution to that of autoradiography, silver-staining detects both strands of the AFLP products giving rise to double-band patterns which are difficult to analyse.

A more reliable approach involves the use of fluorescently-labelled selective primers. Fluorescent amplification products are detected as they migrate through the gel and pass through an on-line laser. Detection of AFLP fragments using an automated sequencer has several advantages. The cost of fluorescently-labelled primers is less than radioactively-labelled primers, images can be obtained within several hours instead of several days, and digital gel images can be rapidly analysed using specialised computer software.

Applications:

To investigate the applicability of AFLP for epidemiological studies of *Salmonella*, different serotypes were screened (Aarts *et al.*, 1998). Reproducible, unique profiles were obtained for each of the 62 serotypes genetically identified by AFLP analysis. Different strains, which were previously identified as being identical, could be distinguished proving that AFLP analysis is suitable for bacterial epidemiology and identification of *Salmonella*. Since this study, AFLP analysis has been used extensively for genotyping different *Salmonella* species and serovars, including *S. typhi* (Nair *et al.*, 2000), *S. enterica* serovar Typhimurium (Lan *et al.*, 2003) and *S. enterica* serovar Sofia (Ross *et*

al., 2003). AFLP analysis was compared with Pulsed-Field Gel Electrophoresis (PFGE) for epidemiological studies of *Salmonella enterica* (Lindstedt *et al.*, 2000). The AFLP technique proved to be fast and reproducible with a discriminatory power equal to that of PFGE.

AFLP was evaluated as an epidemiological typing method for *Eschericia coli* (Arnold *et al.*, 1999), and was applied to the isolates from the EcoR collection. Well-defined and reproducible results were obtained, which correlated closely with the MLEE results of the same isolates. Single Enzyme-AFLP (SE-AFLP) was evaluated as a rapid species- and strain-level identification technique (Velappan *et al.*, 2001). *E. coli, Yersina pestis* and several other pathogenic bacteria were screened. SE-AFLP utilises only the infrequent (6 bp) cutter restriction enzyme and there is no labelled primer in the amplification of the restriction fragments. SE-AFLP generates a lesser number of large fragments, allowing direct visual analysis of fingerprint patterns. Species- and strain-level identification of pathogenic bacteria was achieved by SE-AFLP and this technique is highly reproducible and sensitive.

To determine the taxonomic relationship between the soft rot bacteria *Erwinia carotovora* and *Erwinia chrysanthemi* at the subspecies level, AFLP analysis was performed (Avrova *et al.*, 2002). Subspecies-specific banding profiles were generated which allowed four clusters to be delineated. In addition, the technique identified several unknown isolates and discriminated between closely related strains for epidemiological investigations. AFLP analysis was evaluated for epidemiological and phylogenetic analyses of *Klebsiella* isolates (Jonas *et al.*, 2004). Results allowed the discrimination of *K. pneumoniae* and *K. oxytoca* into two and three phylogenetic groups, respectively. A preliminary cut-off value was also provided, based on AFLP data, for distinguishing epidemiologically non-related *Klebsiella* isolates.

2.5) Conclusions:

Bacterial typing based on variations of the 16S-23S ITS region is simple, rapid and cost-effective. However, this technique may not yield the most reproducible and discriminative results. The most reliable results will be those obtained from amplification of the ITS region of species with multiple rRNA operons (Gürtler and Stanisich, 1996).

Alternatively, AFLP is robust, reproducible and highly discriminative; even more so than rep-PCR which is prone to variations in amplification efficiency. AFLP and rep-PCR were compared with DNA-DNA homology studies, using *Xanthomonas* as a model system (Rademaker *et al.*, 2000). A high correlation was observed from the results, suggesting that AFLP and rep-PCR truly reflect the genotypic and phylogenetic relationships of organisms. Although the initial setup cost for AFLP is expensive, the technique is moderately easy and results can be obtained in as little as two days. A standardised scheme does not yet exist for AFLP as the choice of restriction enzymes, gel systems and visualisation technologies differ from laboratory to laboratory.

Whereas AFLP is initially expensive, PFGE is less so. Studies have proven that the discriminatory power of AFLP is equal to that of PFGE (Lindstedt *et al.*, 2000). However, PFGE is limited in its resolving power which contributes to difficulties with reproducibility and it can take up to four days to obtain analysis results. Nonetheless PFGE is still a viable option for microbial typing of *Pantoea* species, especially for genome mapping.

A RAPID MOLECULAR IDENTIFICATION TECHNIQUE FOR *PANTOEA ANANATIS*

3.1) Introduction:

Pantoea ananatis was isolated for the first time in 1924, from pineapple fruitlets exhibiting brown rot (Serrano, 1928). Since then, *P. ananatis* has been found to cause a wide variety of diseases on a range of agricultural crops including honeydew melon (Wells *et al.*, 1987); cantaloupe fruit (Bruton *et al.*, 1991); onion (Gitaitis and Gay, 1997); sudangrass (Azad *et al.*, 2000) and maize (Paccola-Meirelles *et al.*, 2001). In each of the above studies, the bacterium was characterised based on colony morphology, physiological tests and, in some cases fatty acid analyses. Identification of the causal agent based solely on these methods can prove to be unreliable, as a high similarity in phenotypic characteristics exists, not only between *Pantoea* species, but also among related *Enterobacteriaceae* (Verdonck *et al.*, 1987). In the past, identification based only on phenotypic tests has occasionally led to the incorrect identification of isolates. It is only with the characterisation of the most recent isolates of *P. ananatis*, from onion seed (Walcott *et al.*, 2002), weeds and crops (Gitaitis and Gay, 1997) and *Eucalyptus* (Coutinho *et al.*, 2002), that identification of the bacterium was based on genotypic-, as well as phenotypic- characteristics.

The increasingly frequent occurrences of *P. ananatis* from different hosts has raised many questions concerning the spread of the pathogen, its genetic relatedness to other *Pantoea* species and its presence as an epiphyte on many plants. But most importantly, the regular isolation of *P. ananatis* from these hosts has highlighted the need for a rapid, molecular identification technique for the pathogen. The 16S-23S internally transcribed spacer (ITS) region is variable, flanked and spaced by highly conserved sequences of coding ribosomal genes. There are also multiple copies of these genes in the genome, which facilitates PCR amplification. The 16S-23S ITS region is also more conserved within species than between species, making it a suitable marker for species identification.

Pantoea ananatis was identified from asymptomatic weeds and crops in Georgia, U.S.A, by means of a PCR assay targeted at the 16S-23S internally transcribed spacer (ITS) region (Gitaitis *et al.*, 2002). Species-specific primers (PanITS1 and EC5) were designed, based on the nucleotide sequence of the 16S-23S ITS region of a known *P. ananatis* isolate. An amplicon of 398 bp produced by amplification of DNA with the primers, was representative of a positive identification of *P. ananatis*. The same primers were also used in the detection of *P. ananatis* in naturally infested onion seed (Walcott *et al.*, 2002). In both studies, other *Pantoea* species were tested for specificity and sensitivity towards the designed primers. In both studies, the only other *Pantoea* species to produce a 398 bp amplicon was *P. stewartii* subsp. *stewartii* (Gitaitis *et al.*, 2002, Walcott *et al.*, 2002). In an attempt to prevent the amplification of *P. stewartii* subsp. *stewartii*, Gitaitis *et al.* modified the primers, PanITS1 (5' GTC TGA TAG AAA GAT AAA GAC 3') and EC5 (5' CGG TGG ATG CCC TGG CA 3'), which were used in their original study to PanITS sn2b and PanITS as2b (Personal communication). The modified primers, PanITS sn2b and PanITS as2b, were used in this research project.

The aim of this study, therefore, was to evaluate a PCR-based technique for the rapid identification of *Pantoea ananatis* using primers PanITS sn2b and PanITS as2b, targeting the 16S-23S ITS region.

3.2) Materials and Methods:

3.2.1) 16S-23S ITS-PCR Assay:

Authentic *P. ananatis* strains were received from the U.S.A, Hawaii, Japan, Brazil and South Africa from different hosts. Unidentified isolates, from infected Eucalypts exhibiting *Pantoea*-like symptoms were received from Uganda and three South American countries, hereafter referred to as South America 1, South America 2 and South America 3. Representative strains of *P. agglomerans, P. stewartii* subsp. *stewartii, P. stewartii* subsp. *indologenes, P. dispersa, P. terrea, P. punctata* and *P. citrea* were included in the study together with the type strains of all seven *Pantoea* species.

Nutrient broth (Oxoid) was inoculated with each of the isolates and incubated overnight at 37 ºC. Genomic DNA was extracted using the DNeasy™ tissue kit (QIAGEN) and stored at -20 ºC. DNA was amplified with the species-specific primers PanITS sn2b (5' GTC TGA TAG AAA GAA AAG AAG 3') and PanITS as2b (5' TTC ATA TCA CCT TAC CGG CGC 3') (Inqaba Biotechnologies), targeting the 16S-23S ITS region. Each 50 µL PCR reaction contained 1 x Reaction buffer, 1.5 mM $MgCl_2$, 250 µM of each nucleotide (dATP, dCTP, dGTP, dTTP), 25 pmol of each primer (forward and reverse), 1 U Taq DNA polymerase (Southern Cross Biotechnologies) and 50-100 ng genomic DNA. Amplification was carried out in a GeneAmp 2700 Thermal Cycler (Applied Biosystems). PCR conditions included denaturation at 95 °C for 5 minutes, 30 cycles of denaturation at 95 °C for 1 minute, annealing of primers at 56 °C for 1 minute, primer extension at 72 °C for 1 minute, and final chain elongation at 72 °C for 5 minutes. A negative water control was included with the PCR reactions to monitor contamination. Amplification products were run on a 1.5 % agarose gel (Promega) containing 10 mg/mL ethidium bromide, in 1 x TAE buffer at 100 V for 30 minutes and visualised and photographed under UV light. A 1 000 bp DNA marker (Roche) was run alongside the PCR products.

3.2.2) DNA Sequencing:

The 398 bp products from a small selection of the positive PCR reactions were excised from the agarose gels and purified. Three volumes of Binding buffer (Roche) and one volume of isopropanol were added to the gel slices. After heating at 55 ºC for 10 minutes, the mixture was added to a spin column. The QIAquick™ Purification Kit (QIAGEN) was used to complete the purification.

The purified 16S-23S ITS genes were sequenced with the forward primer, PanITS sn2b. The 10 µL sequencing reaction contained 2 µL Big Dye Sequencing Reaction Mix (ABI Prism), 3.2 pmol forward primer and 150 ng purified template DNA. Sequencing PCR conditions included denaturation at 96 ºC for 5 seconds and 25 cycles of denaturation at 96 ºC for 10 seconds, annealing of primer at 50 ºC for 5 seconds and elongation at 60 ºC

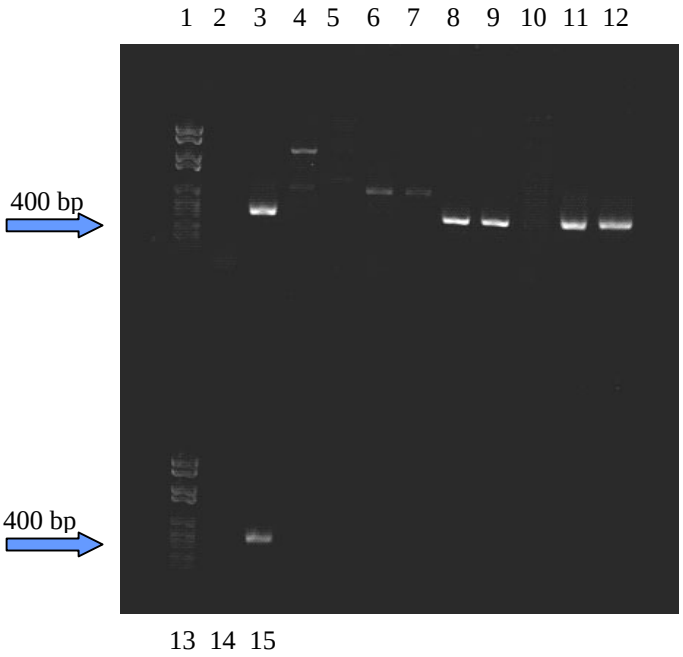
for 4 minutes. PCR products were sequenced on an ABI Prism DNA Automated Sequencer (Perkin-Elmer). Nucleotide sequences were edited and aligned using BioEdit Sequence Alignment Editor v 5.0.9. Heuristic searches with maximum parsimony, using stepwise addition produced a phylogenetic tree in PAUP*4.0b3 (Swofford, 2000). Bootstrap values were obtained from 1 000 bootstrap replicates, using step-wise addition.

3.3) Results:

3.3.1) 16S-23S ITS-PCR Assay:

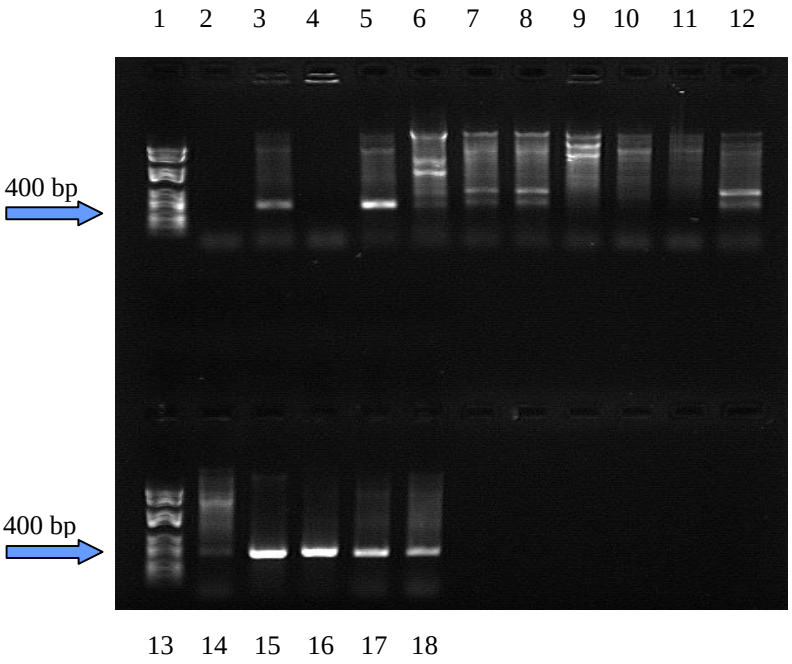
The presence of a 398 bp PCR product was a positive indication of *P. ananatis*, and was visible for all authentic strains (Fig. 3.1). Of the ten unidentified isolates from South America 1, only one (BCC007) yielded a 398 bp product following amplification. The remaining isolates produced different multiple banding patterns or large PCR products of between 600 bp to 900 bp (Fig. 3.1). None of the unidentified isolates from Uganda, South America 2 or 3 produced the correct sized amplicon. Multiple band patterns or smears were visible for the majority of these isolates (Fig. 3.2).

PCR products were not visible for the type strains of *P. agglomerans* (LMG1286^T); *P. stewartii* subsp. *stewartii* (LMG2715^T); *P. dispersa* (LMG2603^T); *P. punctata* (LMG22050^T); *P. terrea* (LMG22051^T) or *P. citrea* (LMG22049^T) (Fig. 3.3). A PCR product of the correct size was visible for the type strain of *P. stewartii* subsp. *indologenes* (LMG2632^T) (Fig. 3.3) and for all representative isolates of this subspecies. A 398 bp band was also present for two isolates received as *P. agglomerans*, from onion bulb (BCC146) (Fig. 3.2) and a *Sorbus* species (LMG2570) (Fig. 3.3), respectively, and possibly for a *P. terrea* isolate (LMG22099) (Fig. 3.3) from persimmon, although the latter PCR product was considerably fainter.



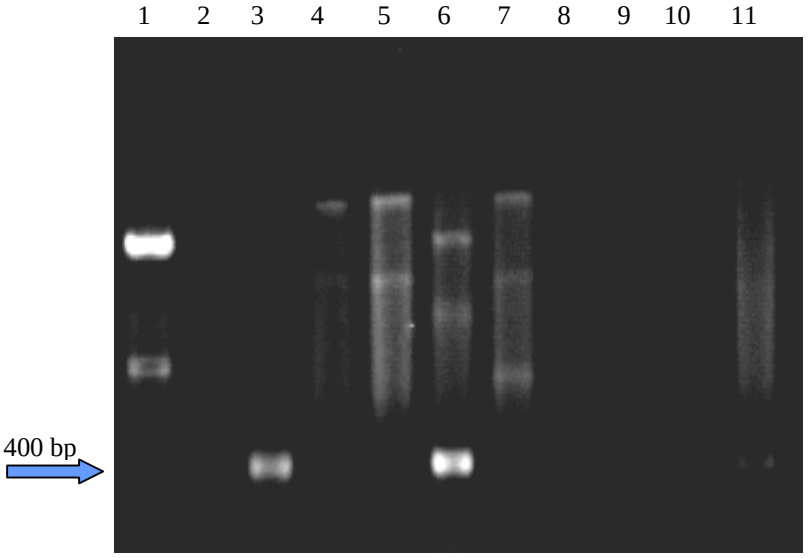
Lane 1: DNA marker VI
Lane 2: Negative control
Lane 3: LMG20103
Lane 4: BCC002
Lane 5: BCC005
Lane 6: BCC004
Lane 7: BCC015
Lane 8: LMG2676
Lane 9: LMG2671
Lane 10: LMG1286^{T}
Lane 11: LMG2665^{T}
Lane 12: LMG2666
Lane 13: DNA marker VI
Lane 14: BCC117
Lane 15: BCC112

Figure 3.1: Agarose gel electrophoresis of the 16S-23S ITS-PCR assay of authentic *Pantoea ananatis* (LMG2665^{T}, LMG20103, LMG2676, LMG2666, BCC112), *Pantoea stewartii* subsp. *indologenes* (LMG2671) and *Pantoea agglomerans* (LMG1286^{T}) strains and possible *P. ananatis* isolates (BCC002, BCC005, BCC004, BCC015, BCC117) from Uganda and South Africa.



Lane 1: DNA marker VI
Lane 2: Negative control
Lane 3: BCC095
Lane 4: BCC077
Lane 5: LMG20104
Lane 6: BCC109
Lane 7: BCC105
Lane 8: BCC107
Lane 9: BCC008
Lane 10: BCC011
Lane 11: BCC003
Lane 12: BCC006
Lane 13: DNA marker VI
Lane 14: BCC145
Lane 15: BCC146
Lane 16: BCC150
Lane 17: BCC154
Lane 18: BCC157

Figure 3.2: Agarose gel electrophoresis of the 16S-23S ITS-PCR assay of authentic *Pantoea ananatis* strains (LMG20104, BCC095, BCC145, BCC146, BCC150, BCC154, BCC157) and possible *P. ananatis* isolates (BCC077, BCC109, BCC105, BCC107, BCC008, BCC011, BCC003, BCC006) from South America and Uganda.



Lane 1: DNA marker XIV
Lane 2: Negative control
Lane 3: LMG2632^{T}
Lane 4: LMG2715^{T}
Lane 5: LMG22049^{T}
Lane 6: LMG2570
Lane 7: LMG22050^{T}
Lane 8: LMG22051^{T}
Lane 9: LMG2603^{T}
Lane 10: LMG22099

Figure 3.3: Agarose gel electrophoresis of the 16S-23S ITS-PCR assay of authentic *Pantoea stewartii* subsp. *indologenes* (LMG2632^{T}), *Pantoea stewartii* subsp. *stewartii* (LMG2715^{T}), *Pantoea agglomerans* (LMG2570), *Pantoea citrea* (LMG22049^{T}), *Pantoea punctata* (LMG22050^{T}), *Pantoea terrea* (LMG22051^{T}, LMG22099) and *Pantoea dispersa* (LMG2603^{T}) strains.

The results from the 16S-23S ITS-PCR assay are summarized in Table 3.1.

Table 3.1: Results from the 16S-23S ITS-PCR assay of authentic *Pantoea* species and possible *P. ananatis* isolates from South America and Uganda. Presence of a 398 bp product is denoted by +, whilst absence of the PCR product is denoted by -.

Code	Species	Location	Host	398 bp
LMG 2665^T	*P. ananatis*	Brazil	Pineapple	+
LMG 2666	*P. ananatis*	Hawaii	Pineapple	+
LMG 2668	*P. ananatis*	Hawaii	Pineapple	+
LMG 20103	*P. ananatis*	Piet Retief	*Eucalyptus*	+
LMG 20104	*P. ananatis*	Harding	*Eucalyptus*	+
LMG 20106	*P. ananatis*	Tzaneen	*Eucalyptus*	+
BCC095	*P. ananatis*	Vryheid	*Eucalyptus*	+
BCC118	*P. ananatis*	KZN	*Eucalyptus*	+
BCC117	?	KZN	*Eucalyptus*	-
BCC112	*P. ananatis*	KZN	*Eucalyptus*	+
BCC114	*P. ananatis*	KZN	*Eucalyptus*	+
BCC154	*P. ananatis*	Limpopo	Onion	+
BCC157	*P. ananatis*	Limpopo	Onion	+
BCC087	*P. ananatis*	USA	Onion	+
BCC083	*P. ananatis*	USA	Onion	+
BCC098	*P. ananatis*	USA	Sudangrass	+
BCC103	*P. ananatis*	USA	Sudangrass	+
BCC099	*P. ananatis*	USA	Sudangrass	+
LMG 2676	*P. ananatis*	USA	Wheat	+
BCC150	*P. ananatis*	USA	Honeydew melon	+
CTB1061	*P. ananatis*	Japan	Rice	+
CTB1135	*P. ananatis*	Japan	Rice	+
LMG 5342	*P. ananatis*	USA	Clinical strain	+
LMG 2678	*P. ananatis*	Zimbabwe	*Puccinia graminis*	+
LMG 2628	*P. ananatis*		Banana	+
LMG 22800	*P. ananatis*	Belgium	Human	+
LMG 1286^T	*P.agglomerans*	Zimbabwe	Human	-
LMG 2565	*P.agglomerans*	Canada	Cereals	-
LMG 2570	*P.agglomerans*	USA	*Sorbus sp.*	+
LMG 2660	*P.agglomerans*	Japan	*Wisteria floribunda*	-
BCC146	*P.agglomerans*	RSA	Onion bulb	+
BCC145	*P. agglomerans*	RSA	Bean leaf	-
LMG 2603^T	*P. dispersa*	Japan	Soil	-
LMG 2602	*P. dispersa*	India	*Sorghum bicolor*	-

LMG 2604	*P. dispersa*	Netherlands	*Rosa sp.*	-
LMG 2749	*P. dispersa*		Human	-
LMG 2632^{T}	*P. stewarti spp. indologenes*	India	Fox millet	+
LMG2631	*P. stewarti spp. indologenes*	India	Millet	+
LMG 2671	*P. stewarti spp. indologenes*	Hawaii	Pineapple	+
LMG 2630	*P.stewartii ssp. indologenes*		Guar gum powder	+
LMG 2673	*P.stewartii ssp. indologenes*	Hawaii	Pineapple	+
LMG 2674	*P.stewartii ssp. indologenes*	Hawaii	Pineapple	+
LMG 2715^{T}	*P. stewartii ssp.stewartii*	USA	Corn	-
LMG 2718	*P. stewartii ssp.stewartii*	USA	Corn	-
LMG 2713	*P. stewartii ssp.stewartii*	USA	Corn	-
LMG 22050^{T}	*P. punctata*	Japan	Mandarin orange	-
LMG 22097	*P. punctata*	Japan	Mandarin orange	-
LMG 22098	*P. punctata*	Japan	Persimmon	-
LMG 22051^{T}	*P. terrea*	Japan	Soil	-
LMG 22100	*P. terrea*	Japan	Persimmon	-
LMG 22099	*P. terrea*	Japan	Persimmon	+
LMG 22049^{T}	*P. citrea*	Japan	Mandarin orange	-
BCC109	?	Uganda	*Eucalyptus*	-
BCC105	?	Uganda	*Eucalyptus*	-
BCC107	?	Uganda	*Eucalyptus*	-
BCC002	?	South America 1	*Eucalyptus*	-
BCC003	?	South America 1	*Eucalyptus*	-
BCC004	?	South America 1	*Eucalyptus*	-
BCC005	?	South America 1	*Eucalyptus*	-
BCC007	?	South America 1	*Eucalyptus*	+
BCC006	?	South America 1	*Eucalyptus*	-
BCC015	?	South America 1	*Eucalyptus*	-
BCC016	?	South America 1	*Eucalyptus*	-
BCC008	?	South America 2	*Eucalyptus*	-
BCC009	?	South America 2	*Eucalyptus*	-
BCC010	?	South America 2	*Eucalyptus*	-
BCC011	?	South America 2	*Eucalyptus*	-
BCC013	?	South America 2	*Eucalyptus*	-
BCC072	?	South America 2	*Eucalyptus*	-
BCC074	?	South America 2	*Eucalyptus*	-

BCC075	?	South America 2	*Eucalyptus*	-
BCC076	?	South America 2	*Eucalyptus*	-
BCC077	?	South America 2	*Eucalyptus*	-
BCC078	?	South America 2	*Eucalyptus*	-
BCC079	?	South America 2	*Eucalyptus*	-
BCC080	?	South America 2	*Eucalyptus*	-
BCC081	?	South America 2	*Eucalyptus*	-
BCC082	?	South America 2	*Eucalyptus*	-
BCC067	?	South America 3	*Eucalyptus*	-

3.3.2) DNA Sequencing:

The most parsimonious phylogenetic tree was selected from sequence analysis of the partial 16S-23S ITS genes of *P. ananatis* and *P. stewartii* subsp. *indologenes* strains. Of the total number of characters used (334), 47 of the variable characters were parsimony-uninformative and 118 were parsimony-informative. The total tree length was 271 steps. The consistency index was 0.841 and the retention index was 0.882.

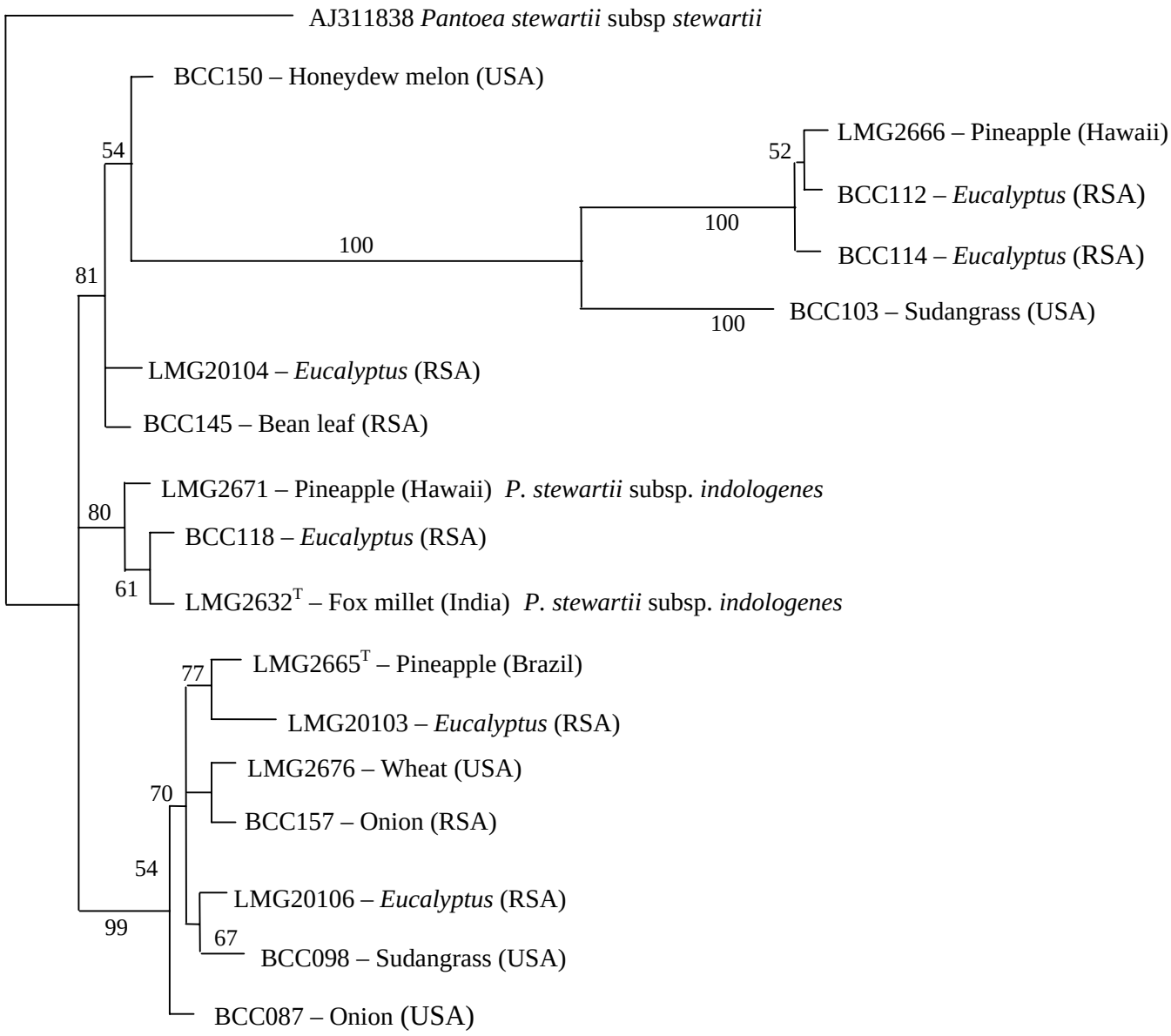


Figure 3.4: Phylogenetic tree of *P. ananatis* and *P. stewartii* subsp. *indologenes* strains based on partial 16S-23S ITS gene sequences (Bootstrap values are indicated on the branches).

3.4) Discussion and Conclusions:

3.4.1) 16S-23S ITS-PCR Assay:

Amplification of a region of the 16S-23S ITS gene with the species-specific primers PanITS sn2b and PanITS as2b has proven to be a rapid, molecular identification technique for *Pantoea ananatis*. However, it was not possible to exclude the non-specific amplification of several isolates previously identified as *P. stewartii* subsp. *indologenes*. The identities of the South American isolates could not be confirmed with the 16S-23S ITS-PCR assay and will be further characterised using alternative methods.

It was expected that authentic strains of *P. ananatis* would produce a PCR product of 398 bp, as the primers were designed based on the ITS sequence of a known *P. ananatis* isolate. The results of the study by Walcott and co-workers state that DNA from *P. stewartii* subsp. *stewartii* also yields a 398 bp product following amplification with their primers (Walcott *et al.*, 2002). Unexpectedly, none of the *P. stewartii* subsp. *stewartii* isolates, including the type strain, produced a 398 bp amplicon following PCR. The type strain and isolates of *P. stewartii* subsp. *indologenes* did, however produce the 398 bp product. Because the original primer sequences were modified it is possible that they now target *P. stewartii* subsp. *indologenes*, instead of *P. stewartii* subsp. *stewartii*, along with *P. ananatis*.

The absence of the correct size PCR product for the Ugandan and South American 2 isolates suggests these isolates are not *P. ananatis*. The fact that only one South American 1 isolate yielded a 398 bp product is unusual. This isolate, BCC007, is assumed to belong to either *P. ananatis* or *P. stewartii* subsp. *indologenes* as these are the only two species detected by the PCR assay. The remaining unidentified isolates from South America 1 may belong to a subspecies of *P. ananatis*, as they produced larger amplicons of 900 bp (BCC002) and 600 bp (BCC004 and BCC015). It is probable that the two isolates received as *P. agglomerans* (BCC146 and LMG2570) were incorrectly

identified, as both yielded the 398 bp amplicon. However, it was not confirmed whether these isolates were *P. ananatis* or *P. stewartii* subsp. *indologenes*.

The presence of a 398 bp PCR product for three out of four unidentified isolates recovered from Eucalypts infected with Coniothyrium canker, suggests that these three isolates (BCC118, BCC112 and BCC114) are *P. ananatis*. It appears that *P. ananatis* can exist in a synergistic relationship with the fungus *Coniothyrium zuluense* on *Eucalyptus* (van Zyl, 1999). Only one isolate recovered from this host did not yield a positive PCR product, BCC117. However, this correlates with previous findings that two *Pantoea* species were originally isolated along with the fungus. 16S rRNA analysis of these two species revealed high similarity to *P. ananatis* and *P. stewartii* subsp. *stewartii*, respectively (van Zyl, 1999).

The three *P. ananatis* isolates from sudangrass (BCC098, BCC099 and BCC103) all produced the correct sized PCR product, as expected. However, isolate BCC103 produced a second, slightly larger band. Despite changes made in annealing temperature and $MgCl_2$ concentration, the larger band was still amplified. This larger band is possibly another copy of the 16S-23S ITS gene containing an insertion, making it of a higher molecular weight.

A rapid, molecular identification technique for *Pantoea ananatis* was evaluated, and found to be species-specific except in the case of *P. stewartii* subsp. *indologenes*. The primer pair PanITS sn2b/PanITS as2b, require further modification and specificity to exclude the amplification of *P. stewartii* subsp. *indologenes* isolates. Further analysis of isolates BCC007, BCC146, LMG2570 and LMG22099 is necessary to determine their correct identification. The identity of the South American isolates remains unknown, but the possibility that they belong to a subspecies of *P. ananatis* cannot be discounted.

3.4.2) DNA Sequencing:

The manner in which the isolates grouped together (Fig. 3.4), provides little information concerning the origin or host specificity of *P. ananatis*. Neither geographical location, nor host plant determined the clustering of the isolates. Therefore at this stage, little can be determined from the phylogenetic tree, except that this pathogen is diverse in its distribution.

The clade containing the type strain of *P. ananatis* (LMG2665^T), and isolates from *Eucalyptus* (LMG20103, LMG20106), wheat (LMG2676), onion (BCC157) and sudangrass (BCC098) reveals that the sequences are closely related. The clustering of two of the unidentified isolates from *Eucalyptus* (BCC112, BCC114) and isolates from pineapple (LMG2666), sudangrass (BCC103) and honeydew melon (BCC150) suggests that these isolates could be a different subspecies of *P. ananatis*. It is possible that three, not two, *Pantoea* species were isolated along with *Coniothyrium zuluense* from *Eucalyptus* in KwaZulu/Natal (van Zyl, 1999). Two of the unidentified Coniothyrium canker isolates (BCC112 and BCC114) group with *P. ananatis* (LMG2666), whilst BCC118 groups with the *P. stewartii* subsp. *indologenes* type strain. The fourth unidentified isolate (BCC117) could belong to *P. stewartii* subsp. *stewartii* as it failed to produce a 398 bp PCR product and therefore was not sequenced.

Due to the fact that only those isolates that yielded a positive PCR product could be sequenced, the resulting dendrogram is not sufficiently informative or significant. The entire 16S-23S ITS region must be amplified and sequenced using universal primers to produce a product that can be examined for polymorphisms between strains and species. In this way, isolates from all seven species of *Pantoea* can be included in the study to possibly determine a pattern in the spread and geographical distribution of plant-pathogenic *Pantoea* isolates.

EVALUATION OF THE ENTIRE 16S-23S ITS REGION FOR MOLECULAR TYPING OF THE GENUS *PANTOEA*

4.1) Introduction:

The genus *Pantoea* was proposed in 1989 to include *Enterobacter agglomerans*, renamed as *Pantoea agglomerans*, and a newly described species, *Pantoea dispersa* (Gavini *et al.*, 1989). Three new species, isolated from soil and fruit samples in Japan, were added to the genus in 1992 as *P. citrea*, *P. punctata* and *P. terrea* (Kageyama *et al.*, 1992). The most recent additions to the genus took place in 1993, when *Erwinia ananas* (synonym *E. uredovora*) and *Erwinia stewartii* were renamed as *Pantoea ananatis* and *Pantoea stewartii* subsp. *stewartii*, respectively; and a new subspecies, *Pantoea stewartii* subsp. *indologenes* was described (Mergaert *et al.*, 1993). The proposal of each of the seven *Pantoea* species was based primarily on DNA hybridisation data, phenotypic tests and electrophoretic protein patterns or quinine composition. To date, seven species and two subspecies of *Pantoea* have been described, yet no standard molecular technique exists for differentiating between *Pantoea* species.

Sequencing of the 16S rRNA gene is widely used for differentiating between bacterial species. Along with DNA-DNA hybridisation, it is considered to be part of the standard description of bacterial taxa. However, 16S rRNA sequences are highly conserved, particularly among the family *Enterobacteriaceae* (Stackebrandt *et al.*, 1999). Many bacteria belonging to this family have genomes containing multiple copies of the rRNA operon. Only minor sequence differences exist between the multiple operons within a bacterium and amongst operons of closely related species (Stackebrandt and Goebel, 1994). Differentiation of *Pantoea* species, based only on 16S rRNA sequence data would not be possible due to this high conservation.

In contrast, the 16S-23S internally transcribed spacer (ITS) region is under less selective pressure than the rRNA structural genes and has exhibited greater sequence variability than the 16S rRNA gene. The size of the ITS region may vary among different species and even among different operons within a single cell, in the case of multiple operons (Condon *et al.*, 1995). Use of the 16S-23S ITS region sequence variability is becoming an important supplement to 16S rRNA sequencing as the standard for differentiating bacterial species (Nagpal *et al.*, 1998).

Differentiation of closely related bacterial species, based on the 16S-23S ITS region, has been achieved by either sequencing of the gene or analysis of band patterns, in the case of polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) or multiple copies of the rRNA operon. Bacterial typing based on spacer-length variation is supposedly rapid, cost-effective and favoured over sequencing (Gürtler and Stanisich, 1996). Sequence analysis of the ITS region has been used to successfully characterise species of the genus *Erwinia*, to which several *Pantoea* species formerly belonged (Duarte *et al.*, 2004; Kim *et al.*, 1999). *P. agglomerans* has been included in studies aimed at identification and detection of Gram-negative bacteria, utilising the 16S-23S ITS region (Jeng *et al.*, 2001; Jensen *et al.*, 1993). Based on the success of these studies, the ITS region seems an appropriate choice for differentiation of *Pantoea* species. Because a partial region of the 16S-23S ITS gene has been examined and deemed too limited for detection of any significant polymorphisms (Chapter 3), it is necessary to examine the entire 16S-23S ITS gene of *Pantoea* species.

The aim of this study was to evaluate a typing technique for the genus *Pantoea*, based on amplification of the entire 16S-23S ITS region using universal primers.

4.2) Materials and Methods:

4.2.1) Amplification of the Entire 16S-23S ITS Region:

Authentic *P. ananatis* strains were received from the U.S.A, Hawaii, Japan, Brazil and South Africa from different plant hosts. Unidentified isolates, from infected Eucalypts exhibiting *Pantoea*-like symptoms were received from Uganda and three South American countries, hereafter referred to as South America 1, South America 2 and South America 3. Representative strains of *P. agglomerans*, *P. stewartii* subsp. *stewartii*, *P. stewartii* subsp. *indologenes*, *P. dispersa*, *P. terrea*, *P. punctata* and *P. citrea* were included in the study together with the type strains of all seven *Pantoea* species.

The 16S-23S ITS gene, from each of the isolates, was amplified with the universal primers FGPS 1490 (5' TGC GGC TGG ATC ACC TCC TT 3') and FGPL 132 (5' CCG GGT TTC CCC ATT CGG 3') (Laguerre *et al.*, 1996) (Inqaba Biotechnologies). Each 50 µL reaction contained 1 x Reaction buffer, 1.5 mM $MgCl_2$, 250 µM of each nucleotide (dATP, dCTP, dGTP, dTTP), 12 pmol of each primer (forward and reverse), 1 U Taq DNA polymerase (Southern Cross Biotechnologies) and 50-100 ng genomic DNA. Amplification was carried out in a GeneAmp 2700 Thermal Cycler (Applied Biosystems). The PCR conditions included denaturation at 95 ºC for 3 minutes, 35 cycles of denaturation at 94 ºC for 30 seconds, annealing of primers at 55 ºC for 30 seconds and primer extension at 72 ºC for 1 minute, and final chain elongation at 72 ºC for 5 minutes. A negative water control was included with the PCR reactions to monitor contamination.

4.2.2) Product Separation:

Agarose Gels:

Amplification products were run on 1.5 % and 3 % agarose gels (Promega) containing 10 mg/mL ethidium bromide, in 1 x TAE buffer at 100 V for 30 minutes. A 1 000 bp DNA marker (Roche) was run alongside the PCR products. The gels were visualised and photographed under UV light.

5 % Polyacrylamide (PAGE) Gels:

Gels were prepared by mixing 5 % acrylamide (BDH Laboratory Supplies), 0.25 % bis-acrylamide (BDH Laboratory Supplies), as a crosslinker and 7.5 % urea in a 50 mM Tris/50 mM Boric acid/1 mM EDTA solution. 10 % ammonium persulphate and 50 µL TEMED were added to the gel solution. Gels were cast using the Bio-Rad gel apparatus and left to polymerise overnight. A 20 minute pre-run was performed at 1 W to equilibrate the ions in the gels and running buffer. PCR samples from the entire 16S-23S ITS-PCR assay, were mixed with formamide loading dye, heated for 5 minutes at 95 ºC and then cooled on ice for 10 minutes. A 2 000 bp marker (Roche) was run alongside the amplification products, at each end of the gel. Electrophoresis was performed at 60 W for 60 minutes, using a 1 x TBE (100 mM Tris/100 mM Boric acid/2 mM EDTA) running buffer. Following electrophoresis, gels were stained in a 1 x SYBR® Gold nucleic acid staining solution (Molecular Probes) for 40 minutes and then visualised and photographed under UV light.

Sequencing Gels:

The forward primer FGPS 1490 was fluorescently labelled with IRD-700 (Biolegio) and amplification was carried out on all isolates as described in section 4.2.1. Sequencing gels were prepared using 20 µL Long Ranger gel stock solution (8 % Long Ranger gel solution (LI-COR Biosciences), 7 M urea, 10 x TBE buffer) and 150 µL 10 % ammonium

persulphate and 15 µL TEMED for polymerisation. Gels were poured using the LI-COR gel casting apparatus and left to polymerise for 60 minutes. A 30 minute pre-run was performed at 1 500 V and 35 W to equilibrate the ions in the gel and running buffer. PCR products were diluted 1:100 with nuclease-free water (Promega), to minimise fluorescent smears, and mixed with an equal volume of formamide loading buffer (95 % formamide, 20mM EDTA, bromophenol blue). The mixture was heated at 90 ºC for 3 minutes and then cooled on ice for 10 minutes. 0.8 µL of each sample was loaded onto the LI-COR gels, along with an IRD-700 labelled sizing standard at each end of the gel. Gels were run on a LI-COR IR^2 automated sequencer (LI-COR Biosciences) for 4 hours at 1 500 V and 42 W with a 0.8 x TBE running buffer. The resulting band patterns were analysed with GelCompar (Applied Maths). Gels were normalised by aligning the 700 bp sizing standards at each end of the gels. Following analysis, a UPGMA dendrogram was constructed using the Dice similarity coefficient. An optimisation setting of 0.09 % and a tolerance setting of 0.48 % was applied to the analysis.

4.3) Results:

4.3.1) Amplification of the Entire 16S-23S ITS Region:

The 16S-23S ITS gene for all authentic strains of *Pantoea* and all unidentified South American and Ugandan isolates was amplified. However, more than one amplification product was visible for each of the isolates included in the study. The multiple amplification products were run on several different types of gels to try and improve the separation and resolution of the bands.

4.3.2) Product Separation:

Agarose Gels:

Amplification products were between 500 and 1 000 bp in size and most were bright and clearly defined when separated on 1.5 % agarose gels (Fig. 4.1). Fainter bands were observed between the brighter bands in some cases (Fig. 4.1, Lane 8). Larger, faint bands of ± 1 000 bp were also present for the majority of isolates. The type strain of *P. ananatis* (LMG2665^T) yielded three amplification products, whilst two PCR products were visible for the type strains of *P. agglomerans* (LMG1286^T) and *P. stewartii* subsp. *indologenes* (LMG2632^T).

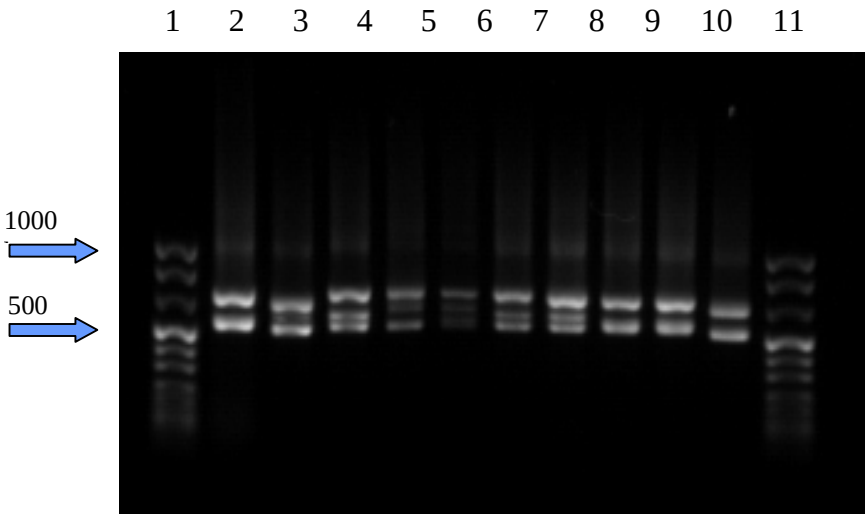


Figure 4.1: Electrophoresis of the entire 16S-23S ITS-PCR assay of authentic *Pantoea ananatis* (BCC095, LMG20106, BCC087, BCC083, LMG2665^T, BCC099), *Pantoea stewartii* subsp. *indologenes* (LMG2632^T) and *Pantoea agglomerans* (LMG1286^T) strains and possible *P. ananatis* isolates (BCC109, BCC003) from South America 1 on a 1.5 % agarose gel

Adjustments to $MgCl_2$ concentration and annealing temperature were made in an attempt to obtain single amplicons. When this failed, the PCR products were run on 3 % agarose gels to improve the resolution. Once separated on a higher percentage gel, there appeared to be between two and five bands for each amplified isolate (Fig. 4.2).

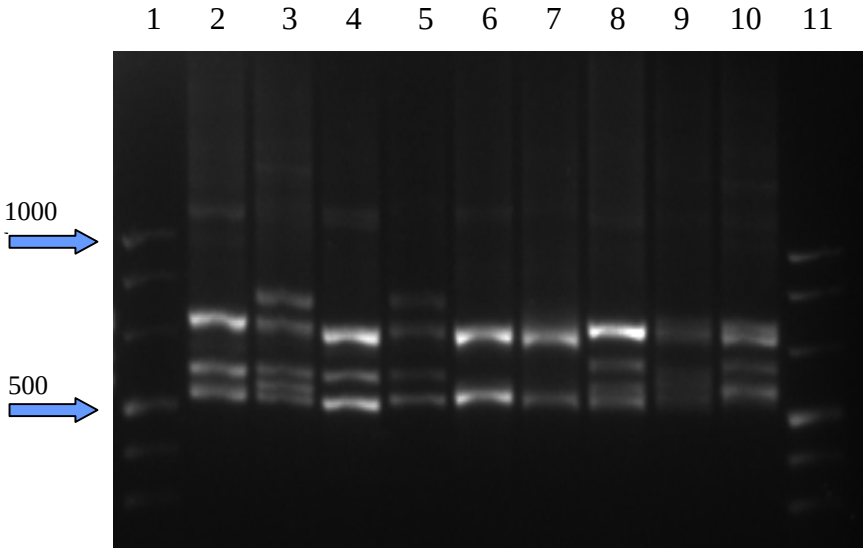


Figure 4.2: Electrophoresis of the entire 16S-23S ITS-PCR of authentic *Pantoea ananatis* (LMG20106, LMG20104, LMG2676, LMG2665^T, LMG20103, LMG2666), *Pantoea stewartii* subsp. *indologenes* (LMG2671, LMG2632^T) and *Pantoea agglomerans* (LMG1286^T) strains on a 3 % agarose gel

5 % Polyacrylamide (PAGE) Gels:

The same PCR products which were run on agarose gels, were separated further on the PAGE gels. The bands ranged from 500 to 2 600 bp in size, but were not clearly defined and most appeared as smears on the gel. The type strain of *P. ananatis* (LMG2665^T), which appeared to produce three to four bands when run on the agarose gels, yielded between six and seven bands when run on a 5 % PAGE gel (Fig. 4.3). The same was true for the type strains of *P. agglomerans* (LMG1286^T) and *P. stewartii* subsp. *stewartii* (LMG2632^T).

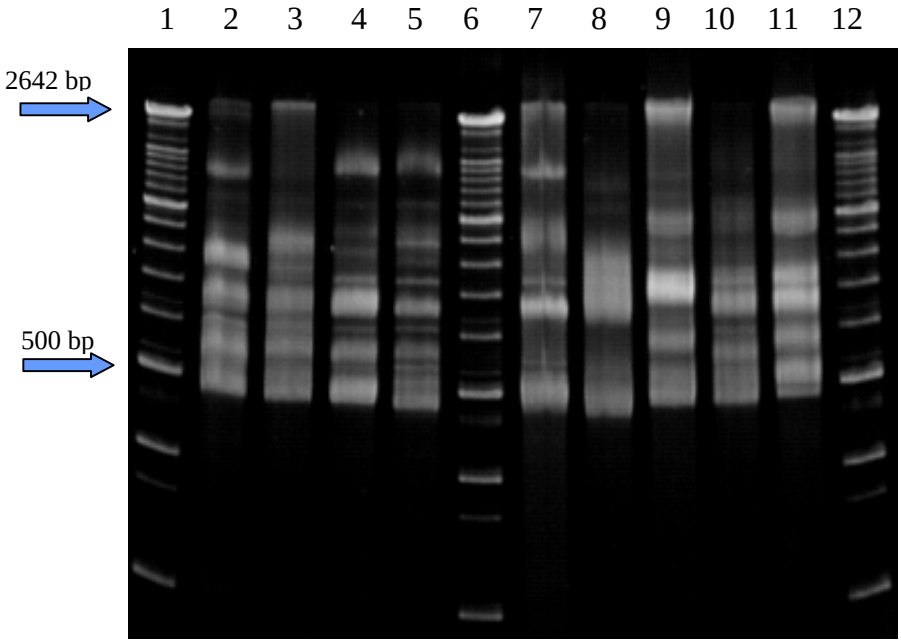


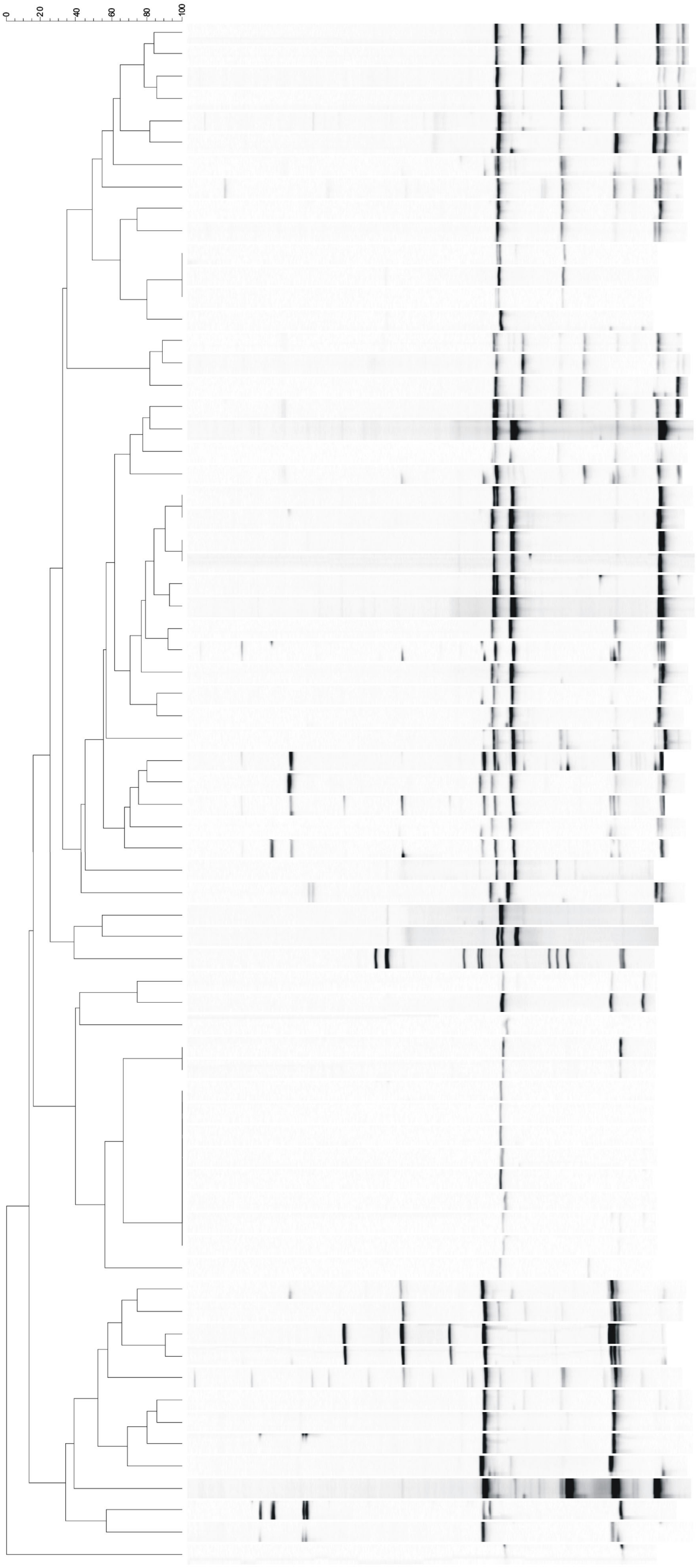
Figure 4.3: Electrophoresis of the entire 16S-23S ITS-PCR of authentic *Pantoea ananatis* (LMG20106, LMG20104, LMG2676, LMG2665^T, LMG20103, LMG2666), *Pantoea stewartii* subsp. *indologenes* (LMG2671, LMG2632^T) and *Pantoea agglomerans* (LMG1286^T) strains on a 5 % PAGE gel

Sequencing Gels:

Clearly defined, multiple amplification products, ranging from 350 to 700+ bp, were visible for all isolates following amplification with a fluorescent primer and separation on an automated sequencing gel. There was a minimum of two bands for each isolate, but not one common-sized band among the isolates. A dendrogram was constructed based on analysis of the band patterns.

Figure 4.4 (on folded insert): UPGMA dendrogram based on band patterns from amplification of the entire 16S-23S ITS regions of species of *Pantoea* and unidentified South American and Ugandan isolates with universal primers FGPS 1490 and FGPL 132. The levels of similarity, representing the Dice similarity co-efficient, are expressed as percentages. The banding patterns adjacent each branch are normalised and background subtracted digitised gel strips processed in GelCompar.

4.4)



Strain	Species			Location	Host
LMG20103	Pantoea	ananatis		Piet Retief	Eucalyptus
LMG2666	Pantoea	ananatis		Hawaii	Pineapple
CTB1061	Pantoea	ananatis		Japan	Rice
CTB1135	Pantoea	ananatis		Japan	Rice
BCC118	Pantoea	ananatis		KZN	Eucalyptus + Con
LMG2632T	Pantoea	stewartii	indologenes	India	Fox millet
LMG20106	Pantoea	ananatis		Tzaneen	Eucalyptus
BCC103	Pantoea	ananatis		USA	Sudangrass
BCC099	Pantoea	ananatis		USA	Sudangrass
LMG2671	Pantoea	stewartii	indologenes	Hawaii	Pineapple
LMG2570	Pantoea	agglomerans		USA	Sorbus sp
LMG2628	Pantoea	ananatis			Banana
LMG2678	Pantoea	ananatis		Zimbabwe	Puccinia graminis
LMG22049T	Pantoea	citrea		Japan	Mandarin orange
BCC083	Pantoea	ananatis		USA	Onion
BCC114	Pantoea	ananatis		KZN	Eucalyptus + Con
LMG2665T	Pantoea	ananatis		Brazil	Pineapple
BCC098	Pantoea	ananatis		USA	Sudangrass
BCC080	?	?		South America 2	Eucalyptus
LMG2676	Pantoea	ananatis		USA	Puccinia graminis
BCC087	Pantoea	ananatis		USA	Onion
BCC077	?	?		South America 2	Eucalyptus
BCC107	?	?		Uganda	Eucalyptus
BCC076	?	?		South America 2	Eucalyptus
BCC082	?	?		South America 2	Eucalyptus
BCC078	?	?		South America 2	Eucalyptus
BCC081	?	?		South America 2	Eucalyptus
BCC015	?	?		South America 1	Eucalyptus
BCC067	?	?		South America 3	Eucalyptus
LMG1286T	Pantoea	agglomerans		Zimbabwe	Human
BCC075	?	?		South America 2	Eucalyptus
BCC105	?	?		Uganda	Eucalyptus
BCC145	Pantoea	agglomerans		RSA	Bean leaf
BCC004	?	?		South America 1	Eucalyptus
BCC008	?	?		South America 2	Eucalyptus
BCC007	?	?		South America 1	Eucalyptus
BCC002	?	?		South America 1	Eucalyptus
BCC013	?	?		South America 2	Eucalyptus
LMG22051T	Pantoea	terrea		Japan	Soil
BCC109	?	?		Uganda	Eucalyptus
LMG22098	Pantoea	punctata		Japan	Persimmon
LMG2660	Pantoea	agglomerans		Japan	Wisteria floribunda
LMG22097	Pantoea	punctata		Japan	Mandarin orange
LMG2604	Pantoea	dispersa		Netherlands	Rosa sp
LMG2718	Pantoea	stewartii	stewartii	USA	Corn
LMG2631	Pantoea	stewartii	indologenes	India	Millet
LMG2674	Pantoea	stewartii	indologenes	Hawaii	Pineapple
LMG2713	Pantoea	stewartii	stewartii	USA	Corn
LMG22050T	Pantoea	punctata		Japan	Mandarin orange
LMG22099	Pantoea	terrea		Japan	Persimmon
LMG22100	Pantoea	terrea		Japan	Persimmon
LMG2602	Pantoea	dispersa		India	Sorghum bicolor
LMG2603T	Pantoea	dispersa		Japan	Soil
LMG2630	Pantoea	stewartii	indologenes		Guar gum powder
LMG2715T	Pantoea	stewartii	stewartii	USA	Corn
LMG2749	Pantoea	dispersa			Human
LMG5342	Pantoea	ananatis		USA	Human
BCC011	?	?		South America 2	Eucalyptus
LMG20104	Pantoea	ananatis		Harding	Eucalyptus
BCC009	?	?		South America 2	Eucalyptus
BCC010	?	?		South America 2	Eucalyptus
BCC095	Pantoea	ananatis		Vryheid	Eucalyptus
BCC150	Pantoea	ananatis		USA	Honeydew melon
BCC154	Pantoea	ananatis		Limpopo	Onion
BCC157	Pantoea	ananatis		Limpopo	Onion
BCC005	?	?		South America 1	Eucalyptus
BCC146	Pantoea	agglomerans		RSA	Onion bulb
BCC072	?	?		South America 2	Eucalyptus
BCC074	?	?		South America 2	Eucalyptus
LMG2673	Pantoea	stewartii	indologenes	Hawaii	Pineapple

4.4) Discussion and Conclusions:

4.4.1) Amplification of the entire 16S-23S ITS Region:

To differentiate among *Pantoea* species, as well as possible *Pantoea ananatis* isolates from South America and Uganda, the entire 16S-23S ITS gene was examined for polymorphisms. The 16S-23S ITS region was amplified for all authentic *Pantoea* strains and unidentified isolates. However, multiple amplification products were observed for all isolates included in the study. These results suggest that the genome of *Pantoea* contains multiple copies of the rRNA operon and therefore, multiple copies of the 16S-23S ITS region.

As universal primers were used to amplify the 16S-23S ITS gene, it was expected that each isolate would yield an amplification product. The presence of multiple amplification products suggests that more than one copy of the 16S-23S ITS region was amplified. This provides a reasonable explanation as *Pantoea* belongs to the family *Enterobacteriaceae* along with *Escherichia coli* and *Salmonella enterica*, both of which contain seven copies of the rRNA operon (Jensen *et al.*, 1993). This was further supported when changes in $MgCl_2$ concentration and an increase in the annealing temperature of the PCR cycle failed to remove the multiple amplicons.

4.4.2) Product Separation:

Agarose Gels:

PCR products were run on a 3 % agarose gel in an attempt to further separate the bands, so a common band could be identified among the isolates. Following electrophoresis, it became clear that a common band among the isolates, notably the type strains of *Pantoea* species, did not exist. It was also not certain if additional amplification products were present for many of the isolates, which would not be visible on a low resolution agarose gel. Poor resolution of band patterns on agarose gels was also observed for

Clostridium difficile (Gürtler, 1993), which was resolved by separating the fragments on long denaturing polyacrylamide gels.

Amplification of the 16S-23S ITS region using universal primers was successful, although more than one amplicon was visible for each isolate. Sequencing or PCR-RFLP analysis of the entire 16S-23S ITS gene was ruled out, as one common band could not be identified among the multiple PCR products. If the resolution of the PCR products could be improved by separation on polyacrylamide gels, it was decided that the band patterns would be treated as DNA fingerprints.

5 % Polyacrylamide (PAGE) Gels:

It was thought that the multiple amplification products would separate further from each other on polyacrylamide gels, as they typically provide a higher resolution than agarose gels. When the polyacrylamide gels were run and stained successfully, the resolution of the PCR products was poor. The 2 000 bp marker separated very well and each size standard was clearly defined. However, the amplification products were smeared and it was impossible to distinguish a clear band pattern. Six to seven bands were visible for the type strain of *P. ananatis* (LMG2665^{T}) and for other *P. ananatis* isolates (LMG20106, LMG20104, LMG2676 and LMG2666) (Fig. 4.2). However the bands did not form a definite pattern when compared to the type strain. Isolate LMG2666 from Hawaii (Lane 11), yielded the band pattern that was most similar to that of the type strain of *P. ananatis* (Lane 9). Although this was not conclusive, as the amplification products were not clearly defined. The resolution of the gels could possibly be improved by using a higher percentage of acrylamide or adjusting the running conditions. However, due to the time-consuming nature of the preparation of polyacrylamide gels, and problems encountered with the staining of the gels, it was decided to discontinue attempts to separate the multiple PCR products on these gels.

Sequencing Gels:

The LI-COR IR^2 automated sequencer provided a rapid and efficient means of separating the multiple amplification products from the entire 16S-23S ITS-PCR assay. Although the cost per gel is relatively high, the technique is simple and the results can be easily interpreted. The amplification products separated well, were clearly defined and very little smearing was visible. Therefore, the multiple PCR products could be treated as DNA fingerprints and were analysed as such.

The UPGMA dendrogram that was constructed, following analysis of the band patterns, revealed little concerning the genetic relatedness of *Pantoea* species and the unidentified isolates. The majority of *P. ananatis* isolates are contained within one large cluster. However, interspersed among the *P. ananatis* isolates are the type strains of *P. stewartii* subsp. *indologenes* (LMG2632^T), *P. citrea* (LMG22049^T) and a *P. agglomerans* isolate (LMG2570). The *P. agglomerans* isolate (LMG2570) produced a positive PCR product when amplified with *P. ananatis* species-specific primers (Chapter 3). The fact that it now groups with *P. ananatis* isolates in the dendrogram with a similarity value of 100 % confirms that LMG2570 was misidentified and is *P. ananatis*, not *P. agglomerans*. The type strain of *P. citrea* (LMG22049^T) was the only isolate of its species to be included in the assay and was expected to cluster with *P. punctata*, as both were isolated from mandarin orange. Instead, LMG22049^T groups with *P. ananatis* isolates with a homology of 80 %.

Several different banding patterns are visible for the *P. ananatis* isolates, although the majority do share at least one common-sized fragment. This intraspecies variability has also been observed for strains of *Staphylococcus aureus*, *Salmonella enterica* and *Escherichia coli* (Pérez-Luz *et al*., 2004). These heterogeneous fragment profiles were correlated with species containing a high copy number of ribosomal operons.

It is unusual that the type strain of *P. stewartii* subsp. *indologenes* (LMG2632^T) falls within the *P. ananatis* cluster with a similarity value of 80 %, especially as the remaining *P. stewartii* subsp. *indologenes* and *P. stewartii* subsp. *stewartii* isolates group together in a cluster with *P. punctata*, *P. terrea* and *P. dispersa*. This suggests that the multiple copies of the rRNA operon are highly similar between certain strains of *Pantoea* species and that the 16S-23S ITS region is not representative of the genotypic differences among *Pantoea* species. This is further illustrated by the large group of *P. punctata*, *P. terrea* and *P. dispersa* isolates which cluster with *P. stewartii* subsp. *indologenes* (LMG2630) and the type strain of *P. stewartii* subsp. *stewartii* (LMG2715^T) with a similarity value of 100 %. All of these isolates seem to possess only one copy of the rRNA operon. Similar behaviour has been observed for *Shigella dysenteriae* and *Shigella sonnei*, both of which contain only two copies of the rRNA operon, where the electrophoretic profiles of the 16S-23S ITS regions of both these species were identical (Pérez-Luz *et al.*, 2004).

It is not possible to distinguish *Pantoea* isolates from each other at or below the species level, based on DNA fingerprints of the 16S-23S ITS region. This is illustrated by the clustering of *P. stewartii* subsp. *stewartii* and *P. stewartii* subsp. *indologenes* with a similarity value of 100 %. The *P. agglomerans* isolates are scattered throughout the dendrogram, as are the unidentified isolates from Uganda and South American countries 1, 2, and 3. A number of the unidentified isolates from South America 1 group with the type strain of *P. agglomerans* (LMG1286^T) with a homology of 80 %, whilst others from South America 1 and 2 cluster with the type strain of *P. terrea* (LMG22051^T) or with a smaller cluster of *P. ananatis* isolates (LMG20104, BCC095, BCC150, BCC154, BCC157).

Although LI-COR automated sequencing gels provided excellent separation of the multiple amplification products, typing using the 16S-23S ITS gene products did not present an accurate portrayal of the genetic relatedness of the genus *Pantoea*. The identity of the Ugandan and South American isolates is still uncertain, as no conclusions can be drawn from the manner in which they clustered in the dendrogram. A high degree of similarity appears to exist among the rRNA operons of species of the genus *Pantoea*,

while heterogeneity was visible between strains of *Pantoea ananatis*. The same behaviour was observed for *Staphylococcus aureus*, where the high level of intraspecies variation was not observed for other *Staphylococcus* species, indicating that it is possible for a single species to produce a variety of spacer amplification products (Jensen *et al.*, 1993). If the genetic relatedness between *Pantoea* species and the unidentified isolates is to be examined successfully, the typing technique will have to be based on screening of the entire genome.

DEVELOPMENT OF AN AFLP-BASED TYPING SYSTEM FOR THE GENUS *PANTOEA*

5.1) Introduction:

Since the first description of *Pantoea ananatis* in 1928 (Serrano, 1928), this bacterium has been subsequently isolated from a range of diverse sources. *P. ananatis* has been identified as the causal agent of bacterial blight and dieback of Eucalypts (Coutinho *et al.*, 2002); leaf spot of maize (Paccola-Meirelles *et al.*, 2001); necrotic leaf blotch disease of sudangrass (Azad *et al.*, 2000); leaf blight, seed stalk rot and bulb decay of onion (Gitaitis and Gay, 1997); postharvest disease of cantaloupe fruit (Bruton *et al.*, 1991) and brown spot of honeydew melon (Wells *et al.*, 1987). Additionally, the presence of the bacterium as an epiphyte on many plant and weed species has been reported (Gitaitis *et al.*, 2002; Coplin and Kado, 2001) and *P. ananatis* has been isolated from a human source (DeBaere *et al.*, 2004). There are also unconfirmed reports of *P. ananatis* causing disease on Eucalypts in South America and Uganda. Despite these regular isolations of *P. ananatis*, little is known concerning the genetic relatedness of strains from different hosts, their host specificity, or the epidemiology of the reported diseases caused by the pathogen.

Understanding the diversity and relationships among pathogenic taxa is an important prerequisite for meaningful taxonomic classification, accurate identification, pathogen detection and epidemiology studies (Avrova *et al.*, 2002). This is particularly important when closely related species are isolated from the same host. *P. ananatis* and *P. agglomerans* have both been reported to cause disease on onion (Gitaitis and Gay, 1997; Hattingh and Walters, 1981); *P. ananatis* and *P. stewartii* subsp. *indologenes* are both thought to cause rot of pineapple (Serrano, 1928; Mergaert *et al.*, 1993) and *P. punctata* and *P. citrea* were both isolated from mandarin orange (Kageyama *et al.*, 1992).

To date, no taxonomic study has been performed on the entire *Pantoea* genus. As a result, it is unclear as to what extent the species of *Pantoea* are genetically related. Also, there is limited 16S rRNA sequence data available in the GenBank/EMBL database for all *Pantoea* species.

Amplified fragment length polymorphism (AFLP) is a genomic fingerprinting method, which has emerged as the technique of choice for taxonomy and genetic diversity studies since its development in 1995 (Vos *et al*., 1995). AFLP can be used for both identification and typing as it can discriminate to below the species level (Savelkoul *et al.*, 1999). AFLP is based on the selective PCR amplification of restriction fragments from a total digestion of genomic DNA with two restriction endonucleases. The AFLP technique involves three steps: digestion of genomic DNA and ligation of double-stranded adaptors to the resulting restriction fragments; selective amplification of restriction fragments with primers complimentary to the ligated adaptors, and gel analysis of the amplified fragments.

AFLP analysis has been used extensively in the identification, classification and epidemiology of bacteria belonging to the *Enterobacteriaceae*, including *Escherichia coli*, *Salmonella*, *Erwinia* and *Klebsiella* (Arnold *et al.*, 1999; Aarts *et al.*, 1998; Avrova *et al.*, 2002; Jonas *et al.*, 2004). Although it is a well-known fact that 16S rRNA sequences are highly conserved, they are still part of the standard description of bacterial species. For this reason, the 16S rRNA genes of representative isolates from South America and Uganda should be sequenced. Therefore the aim of this study was to examine the genetic relatedness of *Pantoea* species and strains, using the AFLP fingerprinting technique and to support the clusters produced by AFLP analysis using 16S rRNA sequencing.

5.2) Materials and Methods:

5.2.1) Amplified Fragment Length Polymorphism (AFLP) Analysis:

Authentic *P. ananatis* strains were received from the U.S.A, Hawaii, Japan, Brazil and South Africa from different hosts. Unidentified isolates, from infected Eucalypts exhibiting *Pantoea*-like symptoms were received from Uganda and three South American countries, hereafter referred to as South America 1, South America 2 and South America 3. Representative strains of *P. agglomerans*, *P. stewartii* subsp. *stewartii*, *P. stewartii* subsp. *indologenes*, *P. dispersa*, *P. terrea*, *P. punctata* and *P. citrea* were included in the study together with the type strains of all seven *Pantoea* species.

Genomic DNA, 50-100 ng, from each of the isolates was digested with 12 U *Eco*RI (Roche) and 8 u *Mse*I (Roche) in 5 x Restriction/Ligation buffer (50 mM Tris HAc, 50 mM MgAc, 250 mM KAc, 25 mM DTT). The digestion reaction was incubated at 37 ºC for 2 hours, then heated at 70 ºC for 15 minutes. Double-stranded adaptors, 5 pmol *Eco*RI and 50 pmol *Mse*I, were added to the 15 µL digestion mixture, together with 5 x Restriction/Ligation buffer, 0.3 mM ATP and 1 U T4 DNA Ligase (Roche). The ligation reaction was incubated at 20 ºC for 2 hours and then diluted 1:10 with nuclease-free water (Promega).

Each 25 µL pre-amplification reaction contained 1 x Reaction buffer, 1.5 mM $MgCl_2$, 250 µM of each nucleotide (dATP, dCTP, dGTP, dTTP), 100 pmol each of Eco-00 (5’ GAC TGC GTA CCA ATT C 3’) and Mse-00 (5’ GAT GAG TCC TGA CTA A 3’) primer (Inqaba Biotechnologies), 1 U Taq DNA polymerase (Southern Cross Biotechnologies) and 2 µL diluted ligation reaction. Amplification was carried out in a GeneAmp Thermal Cycler (Applied Biosystems). The amplification conditions included denaturation at 94 ºC for 3 minutes , 20 cycles of denaturation at 94 ºC for 30 seconds, annealing of primers at 56 ºC for 1 minute and elongation at 72 ºC for 1 minute, and extension at 72 ºC for a further 5 minutes. Following pre-amplification, each reaction was diluted 1:50 with nuclease-free water (Promega).

The selective amplification reaction, in a total volume of 20 µL, contained 1 x Reaction buffer, 1.5 mM $MgCl_2$, 250 µM of each nucleotide (dATP, dCTP, dGTP, dTTP), 0.5 pmol fluorescently-labelled Eco-G primers (5' GAC TGC GTA CCA ATT CG 3'), 2.4 pmol Mse-T primer (5' GAT GAG TCC TGA GTA AT 3') (Inqaba Biotechnologies), 1 U Taq DNA polymerase (Southern Cross Biotechnologies) and 5 µL diluted pre-amplification reaction. Amplification was carried out in a GeneAmp Thermal Cycler (Applied Biosystems). The selective PCR conditions included denaturation at 94 ºC for 5 minutes, 9 cycles of denaturation at 94 ºC for 30 seconds, annealing of primers at 65 ºC for 30 seconds and elongation at 72 ºC for 1 minute, where the annealing temperature decreases by 1 ºC/cycle until 56 ºC is reached. This was followed by 23 cycles of denaturation at 94 ºC for 30 seconds, annealing of primers at 56 ºC for 30 seconds, and elongation at 72 ºC for 1 minute, and a further 5 minutes of extension at 72 ºC.

In an attempt to minimize the number of restriction fragments amplified during the selective amplification, the following additional selective primer combinations were tested: Eco-G/Mse-G, Eco-G/Mse-GC and Eco-G/Mse-CG, where the Eco-G primer was always fluorescently-labelled. The same PCR conditions and primer concentrations were used for all selective amplifications.

5.2.2) LI-COR Gel Analysis:

LI-COR gels were prepared using 20 µL Long Ranger gel stock solution (8 % Long Ranger gel solution (LI-COR Biosciences), 7 M urea, 10 x TBE buffer) and 150 µL 10 % ammonium persulphate and 15 µL TEMED for polymerisation. Gels were poured using the LI-COR gel casting apparatus and left to polymerise for 60 minutes. A 30 minute pre-run was performed at 1 500 V and 35 W to equilibrate the ions in the gel and running buffer. The selective amplification reactions were mixed with an equal volume of formamide loading buffer (95 % formamide, 20mM EDTA, bromophenol blue). The mixture was heated at 90 ºC for 3 minutes and then cooled on ice for 10 minutes. Less than 1 µL of each sample was loaded onto the sequencing gels, along with an IRD-700 labelled sizing standard at each end of the gel. The gels were run on a

LI-COR IR^2 automated sequencer (LI-COR Biosciences) for 4 hours at 1 500 V and 42 W with 0.8 x TBE running buffer. The resulting band patterns were analysed with GelCompar (Applied Maths). Gels were normalised by aligning the 700 bp sizing standards at each end of the gels, and the area between 50-700 bp was analysed. Following analysis, a UPGMA dendrogram was constructed using the Dice similarity coefficient. An optimisation setting of 0.06 % and a tolerance setting of 0.15 % was applied to the analysis.

5.2.3) 16S rRNA Sequencing:

The following possible *Pantoea ananatis* isolates, from South American countries 1, 2, and 3, Hawaii and South Africa were selected as representative strains from several clusters of the AFLP dendrogram for sequencing: BCC002, BCC003, BCC007, BCC009, BCC011, BCC013, BCC067, BCC072, BCC074, BC077, BCC078, BCC079, BCC081, BCC105, BCC117 and LMG2671.

The 16S rRNA gene from each of the selected isolates was amplified using the universal primers 16F27[1] (5' AGA GTT TGA TCC TGG CTC AG 3') and 16R1522[1] (5' AAG GAG GTC ATC CAG CCG CA 3') (Coenye *at al*, 1999). Each 100 µL PCR reaction contained 1 x Reaction buffer, 250 µM of each nucleotide (dATP, dCTP, dGTP, dTTP), 100 pmol of each primer (forward and reverse) (Inqaba Biotechnologies), 1 U Taq Gold (Southern Cross Biotechnologies) and 50-100 ng genomic DNA. Amplification was carried out in a GeneAmp Thermal Cycler (Applied Biosystems). PCR conditions included denaturation at 94 ºC for 5 minutes, 30 cycles of denaturation of 94 ºC for 1 minute, annealing of primers at 58 ºC for 1 minute and elongation at 72 ºC for 1 minute, and a further 5 minutes of extension at 72 ºC. A negative water control was included with the PCR reactions to monitor contamination. The amplification products were loaded onto a 1.5 % agarose gel (Promega) containing 10 mg/mL ethidium bromide and run in 1 x TAE buffer at 100 V for 30 minutes. The 16S rRNA genes were excised from the agarose gels and purified using the QIAquick™ Gel Extraction Kit (QIAGEN).

The 16S rRNA genes were sequenced in both directions using the following internal primers: 16F536 (5' CAG CAG CCG CGG TAA TAC 3'), 16F926 (5' AAC TCA AAG GAA TTG ACG G 3'), 16F1112 (5' AGT CCC GCA ACG AGC GCA AC 3'), 16R339 (5' ACT GCT GCC TCC CGT AGG AG 3'), 16R519 (5' GTA TTA CCG CGG CTG CTG 3'), 16R685 (5' TCT ACG CAT TTC ACC GCT AC 3') and 16R1093 (5' GTT GCG CTC GTT GCG GGA CT 3') (Coenye *et al*., 1999) (Inqaba Biotechnologies). Each 10 µL sequencing reaction contained 2 µL Big Dye Sequencing Reaction Mix (ABI Prism), 1 x Sequencing buffer, 3.2 pmol primer and ± 150 ng purified template DNA. Sequencing amplification conditions included denaturation at 96 ºC for 5 seconds, and 25 cycles of denaturation at 96 ºC for 10 seconds, annealing of primer at 55 ºC and elongation at 60 ºC for 4 minutes. PCR products were sequenced on an ABI Prism DNA Automated Sequencer (Perkin-Elmer). A consensus sequence for each isolate was assembled by manual alignment of the internal sequences using BioEdit Sequence Alignment Editor v 5.0.9. Homology searches were performed on each consensus sequence using the BLAST programme from the GenBank/EMBL database. Sequences of the 16S rRNA gene of several *Pantoea, Erwinia, Escherichia, Shigella, Enterobacter* and *Salmonella* species were obtained from the GenBank/EMBL database and aligned with the consensus sequences in ClustalX (Thompson *et al.*, 1997). A Neighbour-Joining tree was constructed in PAUP*4.0 (Swofford, 2000), following bootstrap analysis from 1 000 replicates.

5.3) Results:

5.3.1) Amplified Fragment Length Polymorphism (AFLP) Analysis and LI-COR Gel Analysis:

The original selective primer combination of Eco-G/Mse-T yielded complex DNA fingerprints. For each of the isolates, a minimum of 120 bands was observed following gel analysis. Generally, between 30 and 50 bands are sufficient for AFLP analysis (Janssen *et al.*, 1996). Therefore, use of the Eco-G/Mse-T primer combination for typing of the genus *Pantoea* was discontinued. The selective primer combinations Eco-G/Mse-G

and Eco-G/Mse-GC both produced band patterns where an excess of 100 bands was visible for each isolate. The use of these primer combinations was also discontinued.

The selective primer combination of Eco-G/Mse-CG yielded well-defined DNA fingerprints, with an average of 90 bands per isolate. When the band patterns generated by the different selective primer combinations for the type strain of *P. ananatis* (LMG2665^T) were compared (Fig. 5.1), it was clear that Eco-G/Mse-CG provided the best resolution. This primer combination was the preferred choice for AFLP analysis of the genus *Pantoea*.

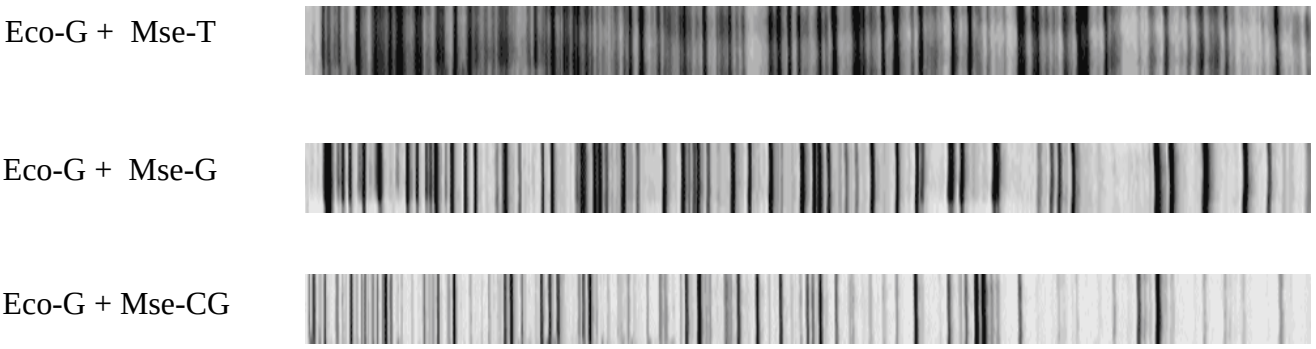
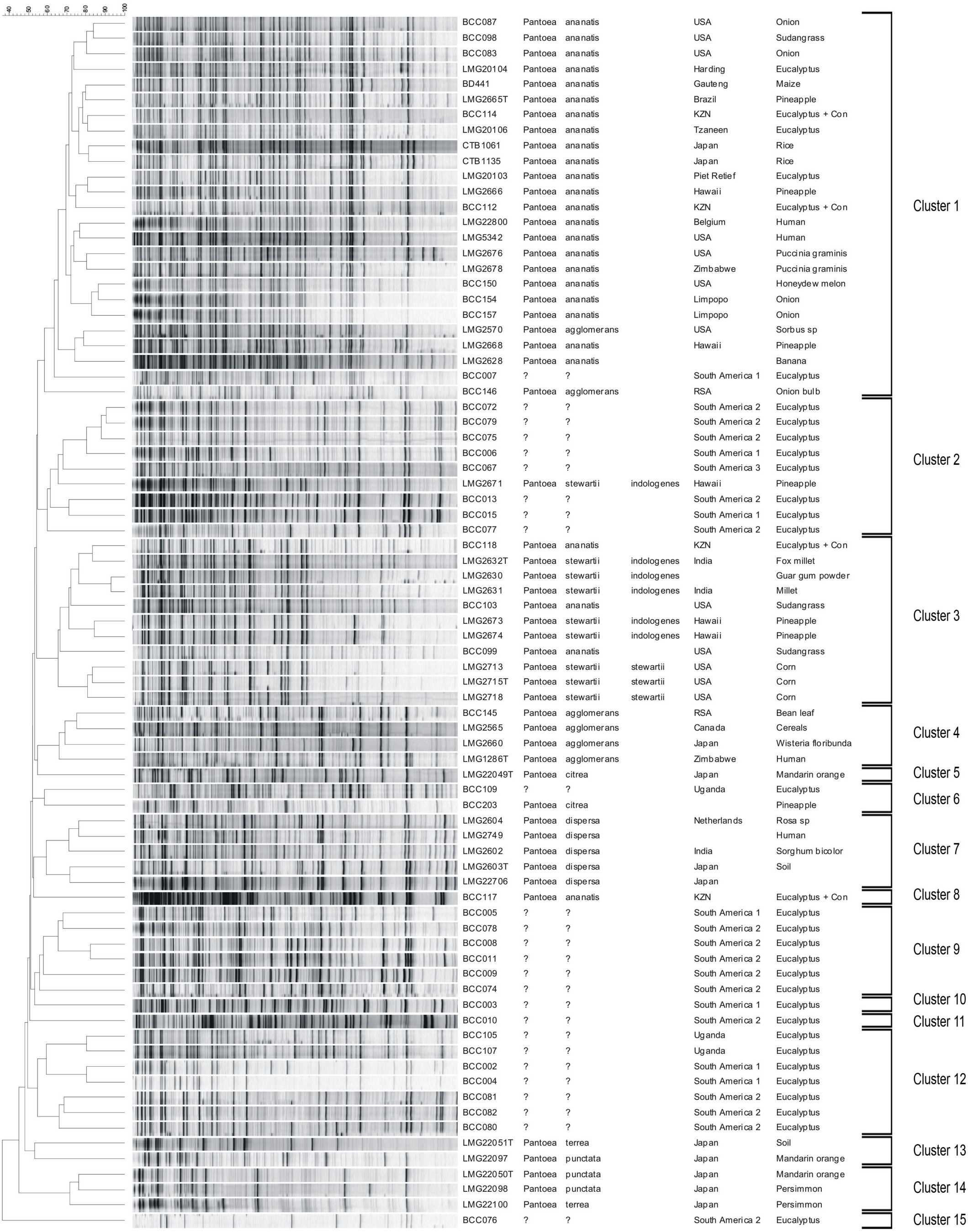


Figure 5.1: AFLP fingerprints of the type strain of *Pantoea ananatis* (LMG2665^T), comparing the band patterns generated by the selective primer combinations Eco-G/Mse-T, Eco-G/Mse-G and Eco-G/Mse-CG.

Figure 5.2 (on folded insert): UPGMA dendrogram based on AFLP analysis of the *Pantoea* species and unidentified South American and Ugandan isolates using the selective primer combination Eco-G/Mse-CG. The levels of similarity representing the Dice similarity co-efficient, are expressed as percentages. The banding patterns adjacent to each branch are normalised and background-subtracted digitised gel strips processed in GelCompar.


40
50
60
70
80
90
100
Strain
Species
Location
Host
BCC087 Pantoea ananatis USA Onion
BCC098 Pantoea ananatis USA Sudangrass
BCC083 Pantoea ananatis USA Onion
LMG20104 Pantoea ananatis Harding Eucalyptus
BD441 Pantoea ananatis Gauteng Maize
LMG2665T Pantoea ananatis Brazil Pineapple
BCC114 Pantoea ananatis KZN Eucalyptus + Con
LMG20106 Pantoea ananatis Tzaneen Eucalyptus
CTB1061 Pantoea ananatis Japan Rice
CTB1135 Pantoea ananatis Japan Rice
LMG20103 Pantoea ananatis Piet Retief Eucalyptus
LMG2666 Pantoea ananatis Hawaii Pineapple
BCC112 Pantoea ananatis KZN Eucalyptus + Con
LMG22800 Pantoea ananatis Belgium Human
LMG5342 Pantoea ananatis USA Human
LMG2676 Pantoea ananatis USA Puccinia graminis
LMG2678 Pantoea ananatis Zimbabwe Puccinia graminis
BCC150 Pantoea ananatis USA Honeydew melon
BCC154 Pantoea ananatis Limpopo Onion
BCC157 Pantoea ananatis Limpopo Onion
LMG2570 Pantoea agglomerans USA Sorbus sp
LMG2668 Pantoea ananatis Hawaii Pineapple
LMG2628 Pantoea ananatis Banana
BCC007 ? ? South America 1 Eucalyptus
BCC146 Pantoea agglomerans RSA Onion bulb
BCC072 ? ? South America 2 Eucalyptus
BCC079 ? ? South America 2 Eucalyptus
BCC075 ? ? South America 2 Eucalyptus
BCC006 ? ? South America 1 Eucalyptus
BCC067 ? ? South America 3 Eucalyptus
LMG2671 Pantoea stewartii indologenes Hawaii Pineapple
BCC013 ? ? South America 2 Eucalyptus
BCC015 ? ? South America 1 Eucalyptus
BCC077 ? ? South America 2 Eucalyptus
BCC118 Pantoea ananatis KZN Eucalyptus + Con
LMG2632T Pantoea stewartii indologenes India Fox millet
LMG2630 Pantoea stewartii indologenes Guar gum powder
LMG2631 Pantoea stewartii indologenes India Millet
BCC103 Pantoea ananatis USA Sudangrass
LMG2673 Pantoea stewartii indologenes Hawaii Pineapple
LMG2674 Pantoea stewartii indologenes Hawaii Pineapple
BCC099 Pantoea ananatis USA Sudangrass
LMG2713 Pantoea stewartii stewartii USA Corn
LMG2715T Pantoea stewartii stewartii USA Corn
LMG2718 Pantoea stewartii stewartii USA Corn
BCC145 Pantoea agglomerans RSA Bean leaf
LMG2565 Pantoea agglomerans Canada Cereals
LMG2660 Pantoea agglomerans Japan Wisteria floribunda
LMG1286T Pantoea agglomerans Zimbabwe Human
LMG22049T Pantoea citrea Japan Mandarin orange
BCC109 ? ? Uganda Eucalyptus
BCC203 Pantoea citrea Pineapple
LMG2604 Pantoea dispersa Netherlands Rosa sp
LMG2749 Pantoea dispersa Human
LMG2602 Pantoea dispersa India Sorghum bicolor
LMG2603T Pantoea dispersa Japan Soil
LMG22706 Pantoea dispersa Japan
BCC117 Pantoea ananatis KZN Eucalyptus + Con
BCC005 ? ? South America 1 Eucalyptus
BCC078 ? ? South America 2 Eucalyptus
BCC008 ? ? South America 2 Eucalyptus
BCC011 ? ? South America 2 Eucalyptus
BCC009 ? ? South America 2 Eucalyptus
BCC074 ? ? South America 2 Eucalyptus
BCC003 ? ? South America 1 Eucalyptus
BCC010 ? ? South America 2 Eucalyptus
BCC105 ? ? Uganda Eucalyptus
BCC107 ? ? Uganda Eucalyptus
BCC002 ? ? South America 1 Eucalyptus
BCC004 ? ? South America 1 Eucalyptus
BCC081 ? ? South America 2 Eucalyptus
BCC082 ? ? South America 2 Eucalyptus
BCC080 ? ? South America 2 Eucalyptus
LMG22051T Pantoea terrea Japan Soil
LMG22097 Pantoea punctata Japan Mandarin orange
LMG22050T Pantoea punctata Japan Mandarin orange
LMG22098 Pantoea punctata Japan Persimmon
LMG22100 Pantoea terrea Japan Persimmon
BCC076 ? ? South America 2 Eucalyptus
Cluster 1
Cluster 2
Cluster 3
Cluster 4
Cluster 5
Cluster 6
Cluster 7
Cluster 8
Cluster 9
Cluster 10
Cluster 11
Cluster 12
Cluster 13
Cluster 14
Cluster 15

In the UPGMA dendrogram, produced from AFLP analysis, 15 distinct clusters are visible. The majority of the *P. ananatis* isolates, including the type strain (LMG2665^T), are contained in cluster 1. These isolates group together with high similarity values of between 70 and 80 %. Two *P. agglomerans* isolates (LMG2570 and BCC146), which both clustered with *P. ananatis* isolates in the 16S-23S ITS dendrogram (Chapter 4), now also group with the *P. ananatis* isolates in the AFLP dendrogram. Only one South American isolate, BCC007, falls within cluster 1.

Cluster 2 includes unidentified isolates from South American countries 1, 2 and 3 and one *P. stewartii* subsp. *indologenes* isolate (LMG2671) from Hawaii. Isolates BCC072, BCC079 and BCC075 from South America 2 cluster together with high similarity values. The only isolate from South America 3 (BCC067), falls within cluster 2 with 66 % homology to BCC072, BCC079 and BCC075.

Both subspecies of *P. stewartii*, *P. stewartii* subsp. *stewartii* and *P. stewartii* subsp. *indologenes*, are contained in cluster 3 and are closely related. Two isolates, received as *P. ananatis* (BCC103 and BCC099) from sudangrass, also fall into cluster 3. The type strain of *P. stewartii* subsp. *indologenes* (LMG2632^T) clusters with BCC118, isolated along with *Coniothryium zuluense* from Eucalypts, with a similarity value of 82 %.

Cluster 4 contains the *P. agglomerans* isolates which group with the type strain LMG1286^T with a similarity value of 66 %. The type strain of *P. citrea* (LMG22049^T) falls singularly in cluster 5, whilst an unidentified Ugandan isolate (BCC109) and an isolate received as *P. citrea* (BCC203) group together in cluster 6 with a homology of 58 %. Cluster 7 includes all of the *P. dispersa* isolates, along with the type strain LMG2603^T. The similarity values in this cluster range from 58 to 72 %. Isolates from South America 2, and one isolate from South America 1 make up cluster 9.

Clusters 8, 10 and 11 each contain only one isolate, respectively BCC117 from KwaZulu/Natal, BCC002 from South America 1 and BCC010 from South America 2. Isolates from Uganda, South America 1 and South America 2 are found in cluster 12. Cluster 13 contains the type strain of *P. terrea* (LMG22051^T) and one *P. punctata* (LMG22097) isolate. These two isolates do not share a similar fingerprint, and are only 54 % homologous. Cluster 14 comprises two *P. punctata* strains, including the type strain (LMG22050^T and LMG22098), and an isolate received as *P. terrea* (LMG22100).

Contained in cluster 15 is an isolate from South America 2 (BCC076). Because BCC076 is linked to the dendrogram with a homology of less than 40 %, it is likely that this isolate does not even belong to the genus *Pantoea*.

5.3.2) 16S rRNA Sequencing:

A region of ±1 500 bp of the 16S rRNA gene was used for each sequenced isolate in the BLAST searches. The results from these homology searches are listed in Table 5.1.

Table 5.1: Results from BLAST searches of the 16S rRNA gene sequences of unidentified isolates from South America, Uganda and South Africa.

Strain		BLAST species	% similarity
From cluster 1:	BCC007	*Pantoea agglomerans*	98 %
From cluster 2:	BCC013	*Pantoea ananatis*	99 %
	BCC067	*Pantoea ananatis*	99 %
	BCC072	*Pantoea agglomerans*	99 %
	BCC077	*Enterobacter* sp.	97 %
	BCC079	*Pantoea agglomerans*	99 %
	LMG2671	*Pantoea stewartii* ssp. *stewartii*	99 %
From cluster 8:	BCC117	*Serratia ficaria*	97 %
From cluster 9:	BCC009	*Escherichia senegalensis*	99 %
	BCC011	*Escherichia senegalensis*	99 %
	BCC074	*Escherichia senegalensis*	99 %
	BCC078	*Escherichia senegalensis*	99 %
From cluster 10:	BCC003	*Enterobacter sakazakii*	97 %
From cluster 12:	BCC002	*Pantoea agglomerans*	99 %
	BCC081	*Pantoea agglomerans*	99 %
	BCC105	*Pantoea ananatis*	99 %

The Neighbour-Joining phylogram was constructed from aligned 16S rRNA gene sequences of South American, Ugandan, South African and Hawaiian isolates and representative strains of related phytopathogenic bacteria and enterobacteria. The majority of South American 2 isolates fall into a cluster with *Escherichia senegalensis*. This grouping is supported with a strong bootstrap value of 93 %. The remaining sequenced South American 2 isolates group in a large cluster with *P. agglomerans* and *P. ananatis* strains, as well as South American 1, South American 3 and Ugandan isolates with a bootstrap value of 93 %. BCC117, from *Eucalyptus* infected with Coniothyrium canker, clusters with *Serratia* species with a bootstrap value of 99 %.

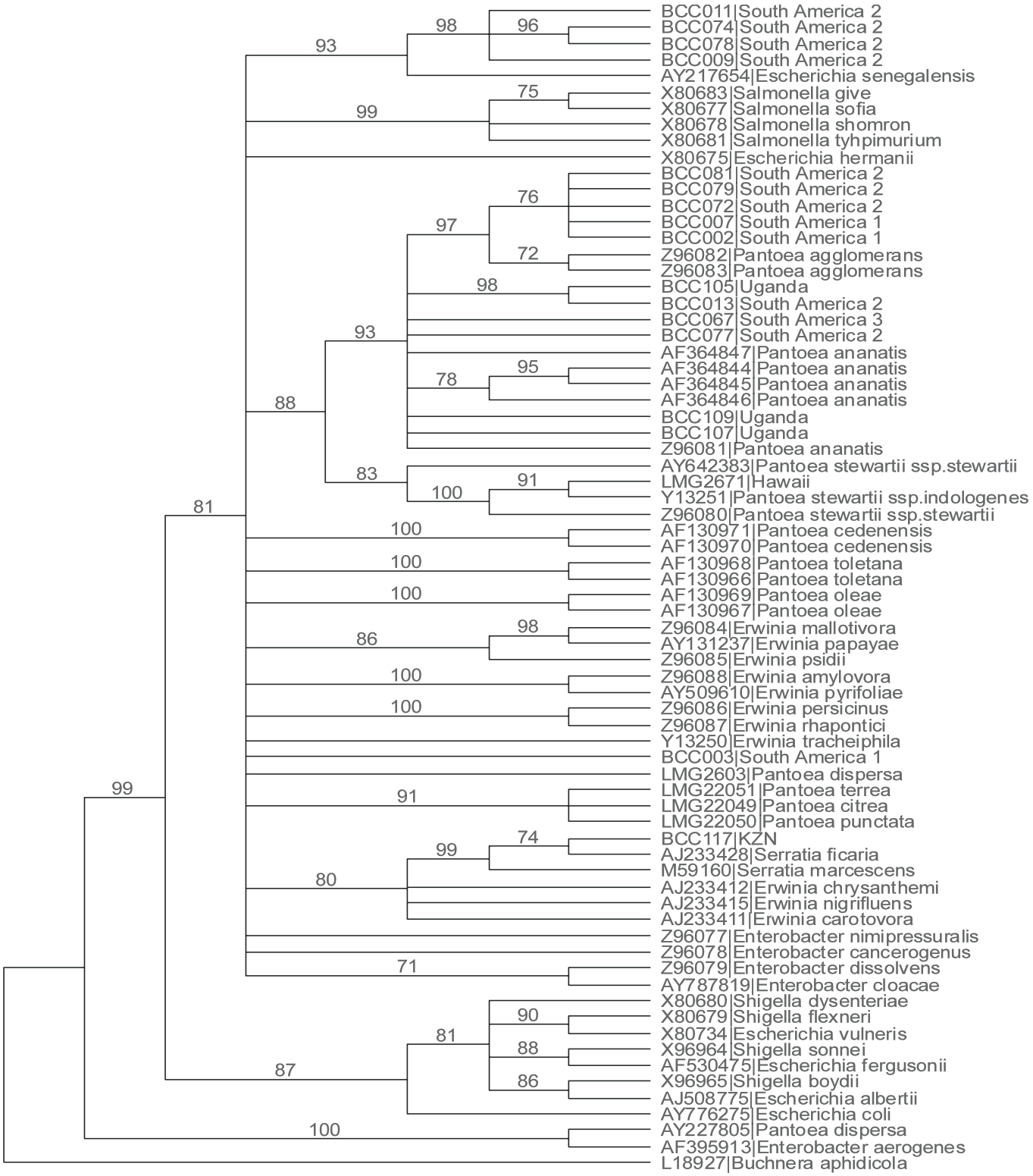


Figure 5.3: Phylogenetic tree of unidentified South American, Ugandan, South African and Hawaiian isolates and related *Enterobacteriaceae* based on 16S rRNA gene sequences (Bootstrap values are indicated on the branches).

5.4) Discussion and Conclusions:

AFLP analysis has proven to be a rapid, reliable technique for typing of the genus *Pantoea*, which yields reproducible results. The DNA fingerprints produced by the selective primer combination Eco-G/Mse-CG were clearly defined, with minimal smearing which did not influence the analysis of the banding patterns. Following gel analysis, 15 distinct clusters could be observed in the UPGMA dendrogram at a cut off point of 54 % (Fig. 5.2). Typing and identification can be standardised by defining windows of similarity, like those specified for the genus *Xanthomonas* (Rademaker *et al.*, 2000). Patterns exhibiting 90-100 % homology are considered to be derived from identical strains, patterns with 60-90 % homology indicate different strains from the same species, whilst 40-60 % homology is obtained with isolates from different species of the same genus. Less than 40 % homology denotes isolates from different genera.

Both LMG2570 and BCC146 produced a positive PCR product when screened with the 16S-23S ITS-PCR assay (Chapter 3) and grouped with *P. ananatis* isolates in the 16S-23S ITS dendrogram (Chapter 4). The appearance of these two isolates in cluster 1 of the AFLP dendrogram (Fig. 5.2) further corroborates their correct identity as *P. ananatis*, not *P. agglomerans*. BCC007 was the only South American isolate to produce a positive PCR product from the 16S-23S ITS-PCR assay (Chapter 3). It was thought that BCC007 belonged to either *P. ananatis* or *P. stewartii* subsp. *indologenes*, as these were the only two *Pantoea* species detected by the assay. The presence of BCC007 in cluster 1, with a similarity value of 62 %, strongly suggests that it belongs to *P. ananatis*. However, the results from the BLAST search of BCC007 reveal 98 % homology to *P. agglomerans* (Table 5.1). Additionally, BCC007 clusters with two *P. agglomerans* strains in the 16S rRNA phylogram (Fig. 5.3).

Cluster 2, of the AFLP dendrogram, is linked to cluster 1 with a similarity value of 56 %, suggesting that this group belongs to either a *P. ananatis* subspecies or a new species, closely related to *P. ananatis*. BCC072 and BCC079 from cluster 2 are both 99 % homologous to *P. agglomerans* and group with this species in the sequencing phylogram,

along with BCC007 and BCC002 from South America 1. This clade is supported by a 76 % bootstrap value. These isolates actually group more closely with *P. ananatis* (cluster 1) in the AFLP dendrogram, with BCC007 falling into that cluster. It is possible that the 16S rRNA genes for *P. ananatis* and *P. agglomerans* strains are so homologous that it is difficult to differentiate between them. This is the case for *Bacillus anthracis* and *Bacillus cereus,* where the 16S rRNA sequence similarity between these two species is 99.9-100 % (Ash *et al.*, 1991). Alternatively, the South American isolates which group closely with *P. agglomerans* in the phylogram could be a new subspecies which can also possibly cause bacterial blight of *Eucalyptus*. BCC013 and BCC067 from cluster 2 are both 99 % homologous to *P. ananatis* and group between *P. agglomerans* and *P. ananatis* in the phylogram.

The appearance of a *P. stewartii* subsp. *indologenes* (LMG2671) isolate in cluster 2 cannot be explained, as the remaining *P. stewartii* strains group together in cluster 3. The BLAST search for LMG2671 revealed 99 % homology to *P. stewartii* subsp. *stewartii* and this isolate forms a clade with both *P. stewartii* subspecies in the phylogram with a bootstrap value of 100 %.

As mentioned, all isolates of both subspecies of *P. stewartii* are contained in cluster 3 of the AFLP dendrogram. BCC103 and BCC099, received as *P. ananatis*, group with *P. stewartii* subsp. *indologenes* in cluster 3. As both isolates gave positive results in the 16S-23S ITS-PCR assay (Chapter 3), there was no reason to doubt their identity as *P. ananatis*. However, because these two isolates group with *P. stewartii* isolates in the AFLP dendrogram, it is probable that they belong to *P. stewartii* subsp. *indologenes*, as the species-specific primers for the 16S-23S ITS-PCR assay detect both *P. ananatis* and *P. stewartii* subsp. *indologenes*. It is possible that BCC103 and BCC099 were incorrectly identified, as both *P. ananatis* and *P. stewartii* were isolated from diseased leaves of sudangrass and there is little phenotypic difference between these two species (Azad *et al.*, 2000).

BCC118 also yielded a positive PCR product when screened with the 16S-23S ITS-PCR assay (Chapter 3) and was thought to be *P. ananatis*. Previous findings, based on 16S rRNA sequencing indicated the presence of two *Pantoea* species isolated with *C. zuluense*, *P. ananatis* and possibly *P. stewartii* subsp. *stewartii* (van Zyl, 1999). Because isolates BCC118, BCC112 and BCC114 were positive for the 16S-23S ITS-PCR assay, and BCC117 was negative, it was assumed that BCC117 was the *P. stewartii* subsp. *stewartii* isolate. However, based on the clustering of these four isolates in the AFLP dendrogram, it is clear that BCC112 and BCC114 belong to *P. ananatis*, as they fall in cluster 1, and BCC118 is actually a *P. stewartii* subsp. *indologenes* strain.

Following the BLAST search, BCC117 was found to be 97 % homologous to *Serratia ficaria*, an enterobacterium involved in the fig tree ecosystem which has also been isolated from human clinical samples (Anahory *et al.*, 1998). BCC117 forms a clade with *Serratia ficaria* and *Serratia marcescens* in the sequencing phylogram. It is not clear if BCC117 was merely present on the leaf at the time of isolation, or if it is a *Serratia* species which can cause disease on *Eucalyptus*. Pathogenicity tests must be performed on this isolate to gauge its pathogenic status.

It was expected that *P. citrea* (BCC203), from cluster 6, would group with the type strain LMG22049^{T} from cluster 5, but as these two isolates share very few common bands, it is probable that BCC203 was mistakenly identified as *P. citrea*. The fingerprint of BCC203 is also dissimilar to that of the Ugandan isolate BCC109, suggesting that these two isolates do not belong to the same species, despite falling into cluster 6 together with a similarity value of 58 %.

No known *Pantoea* species fall into cluster 9 and four strains from this group were selected for 16S rRNA sequencing. All four isolates (BCC009, BCC011, BCC074 and BCC078) showed 99 % homology to *Escherichia senegalensis* following BLAST searches, and group with this strain in the phylogram with a bootstrap value of 93 %. Like BCC117, it is not clear if the South American isolates from cluster 9 in the AFLP dendrogram are pathogenic on *Eucalyptus* and require further examination.

BCC003, from South America 1 singularly forms cluster10 in the AFLP dendrogram and is 97 % similar to *Enterobacter sakazakii*, but groups more closely with *Erwinia* species and *P. dispersa* in the phylogram. *E. sakazakii* is considered to be an opportunistic pathogen and has been isolated from food, the environment and human sources (Lehner *et al*., 2004). The identity of BCC003 is still uncertain, as is its role in causing bacterial blight on *Eucalyptus*.

In cluster 12, two Ugandan isolates (BCC105 and BCC107) form a group with two South American 1 isolates (BCC002 and BCC004). This group is linked to a group of three South American 2 isolates (BCC080, BCC081 and BCC082) with a homology of 52 %. This suggests that cluster 12 could comprise of two new *Pantoea* species, or two subspecies of a new *Pantoea* species. Following the BLAST searches, BCC105 showed 99 % homology to *P. ananatis*, while BCC002 and BCC081, from South America 1 and 2 respectively, are 99 % homologous to *P. agglomerans*. In the phylogram, BCC105 groups with BCC013, BCC067 and BCC077 between *P. agglomerans* and *P. ananatis*, implying that these South American and Ugandan strains could belong to a new *Pantoea* species. DNA-DNA hybridisation must be performed on all unidentified isolates from clusters 1, 2 and 12 of the AFLP dendrogram to confirm if they are indeed a new *Pantoea* species or subspecies of either *P. ananatis* or *P. agglomerans*.

More *P. terrea* strains must be included in the AFLP dendrogram as it appears only the type strain of this species (LMG22051^{T}) from cluster 13, has been correctly identified. The three strains from cluster 14 (LMG22050^{T}, LMG22098 and LMG22100) share many common bands, and have a homology of more than 70 %. It is thought that LMG22100, received as *P. terrea* was incorrectly identified and is actually a *P. punctata* strain.

An AFLP-based typing technique was evaluated for the genus *Pantoea* and found to be successful. Distinct clusters were visible for all seven species of *Pantoea*. The species *P. citrea* and *P. terrea* require further examination due to the unusual grouping of their type strains. Additional *P. citrea* and *P. terrea* isolates must be added to the dendrogram to determine if they would form distinct clusters with their type strains. Based on the clustering of the Ugandan and South American isolates in the phylogram, it is possible that two new species or subspecies of *Pantoea* exist. DNA-DNA hybridisations must be performed to confirm this possibility and to correctly identify the South American and Ugandan isolates.

CONCLUSIONS

Characterisation of microorganisms belonging to the *Enterobacteriaceae* is difficult as a high degree of phenotypic similarity exists in this family. In the past, identification of enterobacteria was based on colony morphology, physiological tests, fatty acid analysis and quinone composition. However, misidentification of isolates often occurred due to the phenotypic resemblance between the enterobacteria. More recently, characterisation of bacteria has been through genotypic methods as well as phenotypic characteristics. Sequencing of the 16S rRNA gene is considered a standard part of the description of bacterial taxa, although it is widely accepted that this gene is highly conserved particularly among the *Enterobacteriaceae*. For this reason, the initial molecular identification technique for *Pantoea ananatis* was based on amplification of the 16S-23S ITS region as it was thought to be more variable.

The 16S-23S ITS region can be used for identification of *P. ananatis* if the species-specific primers are redesigned. When the entire 16S-23S ITS gene was amplified with universal primers, it became clear that the *Pantoea* genome contains multiple copies of the rRNA operon. A high degree of similarity was noted among the number and size of rRNA operons of *Pantoea* species, meaning the genetic relatedness could not be examined using this technique. Possible methods for a taxonomy study of the genus *Pantoea* included both rep-PCR and PFGE, but access to a LI-COR automated sequencer made AFLP an obvious choice. AFLP analysis proved to be a reliable and reproducible method for typing of *Pantoea* species.

The majority of the seven species clustered in distinct groups with their type strains. From the manner in which the South American and Ugandan isolates clustered, it is clear that they belong to either a new *Pantoea* species or subspecies. DNA hybridisations will be performed on these strains to finalise their identity. The conclusions from this research project can be briefly summarised as follows:

- A rapid molecular identification technique for *P. ananatis*, based on the 16S-23S ITS region, is possible providing the species-specific primers are redesigned to exclude amplification of *P. stewartii* subsp. *indologenes* (Chapter 3).

- Examination of a partial region of the 16S-23S ITS gene is not sufficient for the determination of the geographical spread of *P. ananatis* (Chapter 3).

- The genomes of *Pantoea* species contain multiple copies of the rRNA operon, which is in keeping with organisms belonging to the family *Enterobacteriaceae* (Chapter 4).

- It is not possible to determine the genetic relatedness of *Pantoea* species, based on DNA fingerprints produced from amplification of the entire 16S-23S ITS region (Chapter 4).

- A high degree of similarity appears to exist among the rRNA operons of species of the genus *Pantoea* (Chapter 4).

- AFLP analysis provides an efficient method for differentiation of *Pantoea* species, even at the strain level (Chapter 5).

- There is little correlation between results produced by AFLP analysis and 16S rRNA gene sequencing of *Pantoea* species (Chapter 5).

- The possibility of more than one species of *Pantoea* causing disease on *Eucalyptus* should be investigated (Chapter 5).

- DNA-DNA hybridisation must be performed on the South American and Ugandan isolates to determine their position within the genus *Pantoea* (Chapter 5).

- The South American and Ugandan isolates should undergo pathogenicity tests to determine whether they are disease-causing organisms or epiphytes (Chapter 5).

- The species *P. citrea* and *P. terrea* require further examination due to the unusual grouping of their type strains in the AFLP dendrogram (Chapter 5).

Although AFLP analysis provided an accurate portrayal of the genetic relatedness of *Pantoea* species, the epidemiology of *P. ananatis* strains is still unclear. MLST is currently the most discriminative typing method for epidemiological studies and has been used successfully to examine the biodiversity of Vibrios (Thompson *et al*., 2004). If the global epidemiology of *Pantoea ananatis* is to be determined, MLST would provide a viable method.

REFERENCES:

Aarts, H.J.M., van Lith, L.A.J.T., and Keijer, J. 1998. High-resolution genotyping of *Salmonella* strains by AFLP-fingerprinting. Lett. Appl. Microbiol. 26: 131-135

Adhya, S. 1999. Regulation of gene expression: Operons and regulons. Biology of the prokaryotes. J.W. Lengeler, G. Drews, and H.G. Schlegel, eds. Blackwell Science. Georg Thieme Verlag, Stuttgart, Germany. 437-468

Anahory, T., Darbas, H., Ongaro, O., Jean-Pierre, H., and Mion, P. 1998. *Serratia ficaria*: A misidentified or unidentified rare cause of human infections in fig tree culture zones. J. Clin. Microbiol. 36: 3266-3272

Anonymous. 1996. Extract of South African Forestry Facts of the year 1993/1994. Department of Water affairs and Forestry and Forest Owners Association.

Arnold, C., Metherell, L., Willshaw, G., Maggs, A., and Stanley, J. 1999. Predictive fluorescent amplified-fragment length polymorphism analysis of *Escherichia coli*: High-resolution typing method with phylogenetic significance. J. Clin. Microbiol. 37: 1274-1279

Ash, C., Farrow, J., Dorsch, M., Stackebrandt, E., and Collins, M. 1991. Comparative analysis of *Bacillus anthracis, Bacillus cereus*, and related species on the basis of reverse transcriptase sequencing of 16S rRNA. Int. J. Syst. Bacteriol. 41: 343-346

Avrova, A.O., Hyman, L.J., Toth, R.L., and Toth, I.K. 2002. Application of amplified fragment length polymorphism fingerprinting for taxonomy and identification of the soft rot bacteria *Erwinia carotovora* and *Erwinia chrysanthemi*. Appl. Environ. Microbiol. 68: 1499-1508

Azad, H.R., Holmes, G.J., and Cooksey, D.A. 2000. A new leaf blotch disease of sudangrass caused by *Pantoea ananas* and *Pantoea stewartii*. Plant Dis. 84: 973-979

Baudart, J., Lemarchand, K., Brisabois, A., and Lebaron, P. 2000. Diversity of *Salmonella* strains isolated from the aquatic environment as determined by serotyping and amplification of the ribososmal DNA spacer regions. Appl Environ. Microbiol. 66: 1544-1552

Beji, A., Mergaert, J., Gavini, F., Izard, D., Kersters, K., Leclerc, H., and De Ley, J. 1988. 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: 77-88

Ben-Hamouda, T., Foulon, T., Ben-Cheikh-Masmoudi, A., Fendri, C., Belhadj, O., and Ben-Mahrez, K. 2002. Molecular epidemiology of an outbreak multiresistant *Klebsiella pneumoniae* in a Tunisian neonatal ward. J. Med. Microbiol. 52: 427-433

Beyer, P., Al-Babili, S., Ye, X-D., Lucca, P., Schaub, P., Welsch, R., and Potrykus, I. 2002. Golden Rice: Introducing the beta-carotene biosynthesis pathway into rice endosperm by genetic engineering to defeat Vitamin A deficiency. J. Nutr. 132: 506S-510S

Bram, R.J., Young, R.A., and Steitz, J.A. 1980. The ribnucluease III site flanking 23S sequences in the 30S ribosomal precursor RNA of *E. coli*. Cell 19: 393-401

Bruton, B.D., Wells, J.M., Lester, G.E., and Patterson, C.L. 1991. Pathogenicity and characterization of *Erwinia ananas* causing a postharvest disease of cantaloup fruit. Plant Dis. 75: 180-183

Buchan, A., Alber, M., and Hodson, R.E. 2001. Strain-specific differentiation of environmental *Escherichia coli* isolates via denaturing gradient gel electrophoresis (DGGE) analysis of the 16S-23S intergenic spacer region. FEMS Microbiol. Ecol. 35: 313-321

Chalhoub, B.A., Thibault, S., Laucou, V., Rameau, C., Hofte, H., and Cousin, R. 1997. Silver staining and recovery of AFLP amplification products on large denaturing polyacrylamide gels. Bio Techniques 22: 216-218

Chiou, C., Hsu, W., Wei, H., and Chen, J. 2001. Molecular epidemiology of a *Shigella flexneri* outbreak in a mountainous township in Taiwan, Republic of China. J. Clin. Microbiol. 39: 1048-1056

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

Condon, C., Squires, C., and Squires, C.L. 1995. Control of rRNA transcription in *Escherichia coli*. Microbiol. Rev. 59: 623-645

Cooksey, D.A. 1986. Galls of *Gypsophila paniculata* caused by *Erwinia herbicola*. Plant Dis. 70: 464-468

Coplin, D.L., and Kado, C.I. 2001. Laboratory guide for identification of plant pathogenic bacteria: *Pantoea*. N.W. Schaad, J.B. Jones, and W. Chun, eds. 3rd Edition. APS Press, St. Paul, Minnesota, USA. 73-83

Coplin, D.L., Majerczak, D.R., Zhang, Y., Kim, W., Jock, S., and Geider, K. 2002. Identification of *Pantoea stewartii* subsp. *stewartii* by PCR and strain differentiation by PFGE. Plant Dis. 86: 304-311

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

De Baere, T., Verhelst, R., Labit, C., Verschraegen, G., Wauters, G., Claeys, G., and Vaneechoutte, M. 2004. Bacteremic infection with *Pantoea ananatis*. J. Clin. Microbiol. 42: 4393-4395

DeLappe, N., O'Halloran, F., Fanning, S., Corbett-Feeney, G., Cheasty, T., and Cormican, M. 2003. Antimicrobial resistance and genetic diversity of *Shigella sonnei* isolates from western Ireland, an area of low incidence of infection. J. Clin. Microbiol. 41: 1919-1924

Dijkshoorn, L. and Towner, K.J. 2001. An introduction to the generation and analysis of microbial typing data. New approaches for the generation and analysis of microbial typing data. L. Dijkshoorn, K.J. Towner, and M. Struelens, eds. Elsevier Science B.V, Amsterdam, The Netherlands. 1-30

dos Anjos Borges, L.G., Vechia, V.D., and Corção, G. 2003. Characterisation and genetic diversity via REP-PCR of *Escherichia coli* isolates from polluted waters in southern Brazil. FEMS Microbiol. Ecol. 45: 173-180

Duarte, V., De Boer, S.H., Ward, L.J., and de Oliveira, A.M.R. 2004. Characterisation of atypical *Erwinia carotovora* strains causing blackleg of potato in Brazil. J. Appl. Microbiol. 96: 535

Ewing, W.H., and Fife, M.A. 1972. *Enterobacter agglomerans* (Beijerinck) comb. nov. (the herbicola-lathyri bacteria). Int. J. Syst. Bacteriol. 22: 4-11

Fischer, S.G., and Lerman, L.S. 1979. Length-independent separation of DNA restriction fragments in two-dimensional gel electrophoresis. Cell 16: 191-200

García-Martínez, J., Martínez-Murcia, A., Antón, A.I., Rodríguez-Valera, F. 1996[a]. Comparison of the small 16S to 23S intergenic spacer region (ISR) of the rRNA operons of some *Escherichia coli* strains of the ECOR collection and *E. coli* K-12. J. Bacteriol. 178: 6374-6377

García-Martínez, J., Martínez-Murcia, A., Rodríguez-Valera, F, and Zorraquino, A. 1996[b]. Molecular evidence supporting the existence of two major groups in uropathogenic *Escherichia coli*. FEMS Immunol. Med. Microbiol. 14: 231-244

García-Martínez, J., Acinas, S.G., Antón, A.I., and Rodríguez-Valera, F. 1999. Use of the 16S-23S ribosomal genes spacer region in studies of prokaryotic diversity. J. Microbiol. Methods 36: 55-64

García-Martínez, J., Bescós, I., Rodríguez-Sala, J.J., and Rodríguez-Valera, F. 2001. RISSC: a novel database for ribosomal 16S-23S RNA genes spacer regions. Nucl. Acids Res. 29: 178-180

Gavini, F., Mergaert, J., Beji, A., Mielcarek, C., Izard, D., Kerster, K., and 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

Gillings, M., and Holley, M. 1997. Repetitive element PCR fingerprinting (rep-PCR) using enterobacterial repetitive intergenic consensus (ERIC) primers is not necessarily directed at ERIC elements. Lett. Appl. Microbiol. 25: 17-21

Gitaitis, R.D., and Gay, J.D. 1997. First report of leaf blight, seed stalk rot, and bulb decay of onion by *Pantoea ananatis* in Georgia. Plant Dis. 81: 1096

Gitaitis, R., Walcott, R., Culpepper, S., Sanders, H., Zolobowska, L., and Langston, D. 2002. Recovery of Pantoea ananatis, causal agent of centre rot of onion, from weeds and crops in Georgia, USA. Crop protection 21: 983-989

Gitaitis, R.D., Walcott, R.R., Wells, M.L., Diaz Perez, J.C., and Sanders, F.H. 2003. Transmission of *Pantoea ananatis*, causal agent of centre rot of onion, by tobacco thrips, *Frankliniella fusca*. Plant Dis. 87: 675-678

Gürtler, V. 1993. Typing of *Clostridium difficile* strains by PCR-amplification of variable length 16S-23S rDNA spacer regions. J. Gen. Microbiol. 139: 3089-3097

Gürtler, V., and Mayall, B.C. 2001. Genomic approaches to typing, taxonomy and evolution of bacterial isolates. Int. J. Syst. Evol. Microbiol. 51: 3-16

Gürtler, V., and Stanisich, V.A. 1996. New approaches to typing and identification of bacteria using the 16S-23S rDNA spacer region. Microbiology 142: 3-16

Hattingh, M.J., and Walters, D.F. 1981. Stalk and leaf necrosis of onion caused by *Erwinia herbicola*. Plant Dis. 65: 615-618

Hauben, L., Moore, E.R.B., Vauterin, L., Steenackers, M., Mergaert, J., Verdonck, L., and Swings, J. 1998. Phylogenetic position of phytopathogens within the *Enterobacteriaceae*. System. Appl. Microbiol. 21: 384-397

Holt, J.G., Krieg, N.R., Sneath, P.H.A., Staley, J.T., and Williams, S.T. 1994. Bergey's Manual of Determinative Bacteriology. 9th Edition. Williams & Wilkins Company, Baltimore, MD, USA.

Huang, X., Chu, M.C., Engelthaler, D.M., and Lindler, L.E. 2002. Genotyping of a homogeneous group of *Yersinia pestis* strains isolated in the United States. J. Clin. Microbiol. 40: 1164-1173

Hulton, C.S.J., Higgins, C.F., and Sharp, P.M. 1991. ERIC sequences: a novel family of repetitive elements in the genomes of *Escherichia coli*, *Salmonella typhimurium*, and other enterobacteria. Mol. Microbiol. 5: 825-834

Janssen, P., Coopman, R., Huys, G., Swings, J., Bleeker, M., Vos, P., Zabeau, M, and Kersters, K. 1996. Evaluation of the DNA fingerprinting method AFLP as a new tool in bacterial taxonomy. Microbiology 142: 1881-1893

Jeng. R.S., Svircev, A.M., Myers, A.L., Beliaeva, L., Hunter, D.M., and Hubbes, M. 2001. The use of 16S and 16S-23S rDNA to easily detect and differentiate common Gram-negative orchard epiphytes. J. Microbiol. Methods 44: 69-77

Jensen, M.A., and Hubner, R.J. 1996. Use of homoduplex ribosomal DNA spacer amplification products and heteroduplex cross-hybridisation products in the identification of *Salmonella* serovars. Appl. Environ. Microbiol. 62: 2741-2746

Jensen, M.A., Webster, J.A., and Straus, N. 1993. Rapid identification of bacteria on the basis of polymerase chain reaction-amplified ribosomal DNA spacer polymorphisms. Appl. Environ. Microbiol. 59: 945-954

Jonas, D., Spitzmüller, B., Daschner, F.D., Verhoef, J., and Brisse, S. 2004. Discrimination of *Klebsiella pneumoniae* and *Klebsiella oxytoca* phylogenetic groups and other *Klebsiella* species by use of amplified fragment length polymorphism. Res. Microbiol. 155: 17-23

Kageyama, B., Nakae, M., Yagi, S., and Sonoyama, T. 1992. *Pantoea punctata* sp. nov., *Pantoea citrea* sp. nov., and *Pantoea terrea* sp. nov. isolated from fruit and soil samples. Int. J. Syst. Bacteriol. 42: 203-210

Kim, H.M., and Song, W.Y. 1996. Characterization of ribosomal RNA intergenic spacer region of several seedborne bacterial pathogens of rice. Seed. Sci. Technol. 24: 571-580

Kim, W.S., Gardan, L., Rhim, S.L., and Geider, K. 1999. *Erwinia pyrifoliae* sp. nov., a novel pathogen that affects Asian pear trees (*Pyrus pyrifolia* Nakai). Int. J. Syst. Bacteriol. 49: 1952

Kwok, S., Kellog, D.E., McKinney, N., Spasic, D., Goda, L., Levenson, C., and Sninsky, J.J. 1990. Effects of primer-template mismatches on the polymerase chain reaction: Human immuno-deficiency virus 1 model studies. Nucleic Acids Res. 18: 999-1005

Lan, R., Davison, A.M., Reeves, P.R., and Ward, L.R. 2003. AFLP analysis of *Salmonella enterica* serovar Typhimurium isolates of phage types DT 9 and DT 135: Diversity within phage types and its epidemiological significance. Microbes Infect. 5: 841-850

Laguerre, G., Mavingui, P., Allard, M., Charnay, M., Louvrier, P., Mazurier, S., Bigottier-Gois, L., and Amarger, N. 1996. Typing of Rhizobia by PCR DNA fingerprinting and PCR-restriction fragment length polymorphism analysis of chromosomal and symbiotic gene regions: Application to *Rhizobium leguminosarum* and its different biovars. Appl. Environ. Microbiol. 62: 2029-2036

Lehner, A., Tasara T., and Stephan, R. 2004. 16S rRNA gene based analysis of *Enterobacter sakazakii* strains from different sources and development of a PCR assay for identification. BCM Microbiol. 4: 43

Lindow, S.E. 1983. The role of bacterial ice nucleation in frost injury to plants. Annu. Rev. Phytopathol. 21: 363-384

Lindstedt, B., Heir, E., Vardund, T., and Kapperud, G. 2000. Fluorescent amplified-fragment length polymorphism genotyping of *Salmonella enterica* subsp. *enterica* serovars and comparison with pulsed-field gel electrophoresis typing. J. Clin. Microbiol. 38: 1623-1627

Louws, F.J., Rademaker, J.L.W., and de Bruijn, F.J. 1999. The three D's of PCR-based genomic analysis of phytobacteria: Diversity, detection, and disease diagnosis. Annu. Rev. Phytopathol. 37: 81-125

Martin, B., Humber, O., Camara, M., Guenzi, E., Walker, J., Mitchell, T., Andrew, P., Prudhomme, M., Alloing, G., and Hakenbeck, R. 1992. A highly conserved repeated DNA element located in the chromosome of *Streptococcus pneumoniae*. Nucleic Acids Res. 20: 3479-3483

McGhee, G.C., Schnabel, E.L., Maxson-Stein, K., Jones, B., Stromberg, V.K., Lacy, G.H., and Jones, A.L. 2002. Relatedness of chromosomal and plasmid DNAs of *Erwinia pyrifoliae* and *Erwinia amylovora*. Appl. Environ. Microbiol. 68: 6182-6192

McGuire, G., Wright, F., and Prentice, M.J. 1997. A graphical method for detecting recombination in phylogenetic data sets. Mol. Biol. Evol. 14: 1125-1131

Mergaert, J., Verdonck, L., Kersters, K., Swings, J., Boeufgras, J., and De Ley, J. 1984. Numerical taxonomy of *Erwinia* species using API systems. J. Gen. Microbiol. 130: 1893-1810

Mergaert, J., Verdonck, L., and 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

Muniruzzaman, S., Tokunaga, H., and Izumori, K. 1994. Isolation of *Enterobacter agglomerans* strain 221e from soil, a potent D-tagatose producer from galactitol. J. Ferment. Bioeng. 78: 145-148

Nagpal, M.L., Fox, K.F., and Fox, A. 1998. Utility of 16S-23S rRNA spacer region methodology: How similar are interspace regions within a genome and between strains for closely related organisms? J. Microbiol. Methods 33: 211-219

Nair, S., Schreiber, E., Thong, K., Pang, T., and Altwegg, M. 2000. Genotypic characterization of *Salmonella typhi* by amplified fragment length polymorphism fingerprinting provides increased discrimination as compared to pulsed-gel electrophoresis and ribotyping. J. Micrbiol. Methods 41: 35-43

Nizan-Koren, R., Manulis, S., Mor, H., Iraki, N.M., and Barash, I. 2003. The regulatory cascade that activates the Hrp regulon in *Erwinia herbicola* pv. *gypsophilae*. Mol. Plant Microbe Interact. 16: 249-260

Norman, D.J., Yuen, J.M.F., Resendiz, R., and Boswell, J. 2003. Characterization of *Erwinia* populations from nursery retention ponds and lakes infecting ornamental plants in Florida. Plant Dis. 87: 193-196

Olive, D.M.,. and Bean, P. 1999. Principles and applications of methods for DNA-based typing of microbial organisms. J. Clin. Microbiol. 37: 1661-1669

Owen, R.J. 2004. Bacterial Taxonomics: Finding the wood through the phylogenetic trees. Methods Mol. Biol. 266: 353-384

Paccola-Meirelles, L.D., Ferreira, A.S., Meirelles, W.F., Marriel, I.E., and Casela, C.R. 2001. Detection of a bacterium associated with a leaf spot disease of maize in brazil. J. Phytopathol. 149: 275-279

Pérez-Luz, S., Yáñez, M.A., and Catalán, V. 2004. Identification of waterborne bacteria by the analysis of 16S-23S rRNA intergenic spacer region. J. Appl. Microbiol. 97: 191-204

Pinto, N.F.J.A. 1995. Produção de picnidios in vivo e in vitro em folhas de milho tratadas com fungicidas. Fitopatologia Brasileira 20: 333

Pujol, C., and Kado, C.I. 2000. Genetic and biochemical characterization of the pathway in *Pantoea citrea* leading to pink disease of pineapple. J. Bacteriol. 182: 2230-2237

Rademaker, J.L.W., Hoste, B., Louws, F.J., Kersters, K., Swings, J., Vauterin, L, Vauterin, P., and de Bruijn, F.J. 2000. Comparison of AFLP and rep-PCR genomic fingerprinting with DNA-DNA homology studies: Xanthomonas as a model system. Int. J. Syst. Evol. Microbiol. 50: 665-677

Ross, I.L., Willmore, R., and Heuzenroeder, M.W. 2003. A fluorescent amplified fragment length polymorphism study of *Salmonella enterica* serovar Sofia, the major *Salmonella* serovar isolated from chickens in Australia. Int. J. Med. Microbiol. 293: 371

Savelkoul, P.H.M., Aarts, H.J.M., De Haas, J., Dijkshoorn, L., Buim, B., Otsen, M., Rademaker, J.L.W., Schouls, L., and Lenstra, J.A. 1999. Amplified-fragment length polymorphism analysis: The state of an art. J. Clin. Microbiol. 37: 3083-3091

Serrano, F.B. 1928. Bacterial fruitlet brown rot of pineapples in the Philippines. Philipp. J. Sci. 36: 271-305

Shah, S.A., and Romick, T.L. 1997. Subspecies differentiation of *Salmonella* by PCR-RFLP of the ribosomal operon using universal primers. Letters Appl. Microbiol. 25: 54-57

Shurtleff, M.C., and Averre, C.W. III. 1997. Glossary of Plant-Pathological Terms. St. Paul, MN: Am. Phytopathol. Assoc.

Sigee, D.C. 1993. Bacterial plant pathology: Cell and molecular aspects. Cambridge University Press, Cambridge, U.K

Stackebrandt, E., and Goebel, B.M. 1994. Taxonomic note: A place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. Int. J. Syst. Bacteriol. 44: 846-849

Stackebrandt, E., Tindall, B., Ludwig, W., and Goodfellow, M. 1999. Prokaryotic diversity and systematics. Biology of the prokaryotes. J.W. Lengeler, G. Drews, and H.G. Schlegel, eds. Blackwell Science. Georg Thieme Verlag, Stuttgart, Germany. 674-722

Stern, M.J., Ames, G.F.L., Smith, N.H., Robinson, E.C., and Higgins, C.F. 1984. Repetitive extragenic palindromic sequences: a major component of the bacterial genome. Cell 37: 1015-1026

Struelens, M.J., De Ryck, R., and Deplano, A. 2001. Analysis of microbial genomic macrorestriction patterns by pulsed-field gel electrophoresis (PFGE) typing. New approaches for the generation and analysis of microbial typing data. L. Dijkshoorn, K.J. Towner, and M. Struelens, eds. Elsevier Science B.V, Amsterdam, The Netherlands. 159-176

Swofford, D.L. 2000. PAUP*: Phylogenetic Analysis Using Parsimony and Other Methods (software). Sinauer Associates, Sunderland, MA

Talon, D., Menget, P., Thouverez, M., Thiriez, G., Haore, H.G., Fromentin, C., Muller, A., and Bertrand, X. 2004. Emergence of *Enterobacter cloacae* as a common pathogen in neonatal nits: Pulsed-field gel electrophoresis analysis. J. Hosp. Infect. 57: 119-125

Tenover, F.C., Arbeit, R.D., Goering, R.V., Mickelsen, P.A., Murray, B.E., Persing, D.H., and Swaminathan, B. 1995. Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: Criteria for bacterial strain typing. J. Clin. Microbiol. 33: 2233-2239

Thompson, F.L., Iida, T., and Swings, J. 2004. Biodiversity of Vibrios. Microbiol. Mol. Biol. Rev. 68: 403-431

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. 1997. The ClustalX-Windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucl. Acids Res. 25: 4876-4882

Towner, K.J., and Grundmann, H. 2001. Generation and analysis of RAPD fingerprinting profiles. New approaches for the generation and analysis of microbial typing data. L. Dijkshoorn, K.J. Towner, and M. Struelens, eds. Elsevier Science B.V, Amsterdam, The Netherlands. 135-157

Vandamme, P., Pot, B., Gillis, M., de Vos, P., Kersters, K., and Swings, J. 1996. Polyphasic taxonomy, a consensus approach to bacterial systematics. Microbiol. Rev. 60: 407-438

Van Zyl, L.M. 1999. Factors associated with Coniothyrium canker of *Eucalyptus* in South Africa. PhD thesis. University of the Orange Free State, Bloemfontein, South Africa.

Velappan, N., Snodgrass, J.L., Hakovirta, J.R., Marrone, B.L., and Burde, S. 2001. Rapid identification of pathogenic bacteria by single-enzyme amplified fragment length polymorphism analysis. Diag. Microbiol. Infect. Dis. 39: 77-83

Verdonck, L., Mergaert, J., Rijckaert, C., Swings, J., Kersters, K., and De Ley, J. 1987. The genus *Erwinia*: A numerical analysis of phenotypic features. Int. J. Syst. Bacteriol. 37: 4-18

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., and Zabeau, M. 1995. AFLP: A new technique for DNA fingerprinting. Nucl. Acids. Res. 23: 4407-4414

Walcott, R.R., Gitaitis, R.D., Castro, A.C., Sanders, F.H., Jr., and Diaz-Perez, J.C. 2002. Natural infestation of onion seed by *Pantoea ananatis*, causal agent of center rot.
Plant Dis. 86: 106-111

Watanabe, M., and Arai, S. 1987. Freezing of water in the presence of ice nucleation active bacterium, *Erwinia ananas*, and its application for efficient freeze-drying of foods. Agric. Biol. Chem. 51: 557-563

Watanabe, K., Kawakita, H., and Sato, M. 1996. Epiphytic bacterium, *Erwinia ananas*, commonly isolated from rice plants and brown planthoppers (*Nilaparvata lugens*) in hopperburn patches. Appl Entomol. Zool. 31: 459-462

Watson, J.D., Hopkins, N.H., Roberts, J.W., Steitz, J.A., and Weiner, A.M. 1987. Molecular biology of the gene. Menlo Park, CA: The Benjamin/Cummings Publishing Company

Wayne, L.G., Brenner, D.J., Colwell, R.R., and 9 other authors. 1987. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. Int. J. Syst. Bacteriol. 37: 463-464

Wells, J.M., Sheug, W.S., Ceponis, M.J., and Chen, T.A. 1987. Isolation and characterization of strains of *Erwinia ananas* from honeydew melons. Phytopathology 77: 511-514

Weigel, R.M., Qiao, B., Teferedegne, B., Suh, D.K., Barber, D.A., Isaacson, R.E., and White, B.A. 2004. Comparison of pulsed field gel electrophoresis and repetitive sequence polymerase chain reaction as genotyping methods for detection of genetic diversity and inferring transmission of *Salmonella*. Vet. Microbiol. 100: 205-217

Wilson, M.P., and Spencer, R.C. 1999. Laboratory role in the management of hospital acquired infections. J. Hosp. Infect. 42: 1-6

Winslow, C.E.A., Broadhurst, J., Buchanan, R.E., Krumwiede, C. Jr., Rogers, L.A., and Smith, G.H. 1920. The families and genera of the bacteria. Final report of the Committee of the Society of American Bacteriologists on characterization and classification of bacterial types. J. Bacteriol. 5: 191-229

Xie, G. 2001. First report of palea browning in China and characterization of the causal organism by phenotypic tests and Biolog. Int. Rice Res. Notes: 25-26

Zabeau, M., and Vos, P. 1993. Selective restriction fragment amplification: A general method for DNA fingerprinting. European Patent Office, publication 0 534 858 A1

Zasypkin, D.V., and Lee, T-C. 1999. Extracellular ice nucleators from *Pantoea ananas*: Effect of freezing on model foods. J. Food Sci. 64: 473-478

Zhang, Q., Weyant, R., Steigerwalt, A.G., White, L.A., Melcher, U., Bruton, B.D., Pair, S.D., Mitchell, F.L., and Fletcher, J. 2003. Genotyping of *Serratia marcescens* strains associated with cucurbit yellow vine disease by repetitive elements-based polymerase chain reaction and DNA-DNA hybridisation. Phytopathology 93: 1240-1246