Enhanced phylogenetic analysis and targeted search for the genus *Kribbella*

Sarah Maureen Curtis

Thesis presented for the Degree of
Doctor of Philosophy
in the Department of Molecular and Cell Biology,
Faculty of Science
University of Cape Town

February 2015

Table of Contents

Acknowledgements

I would like to express my deepest thanks to my supervisor Dr. Paul Meyers for his dedication, time and passion for microbiology. His constant support was essential for the successful completion of this thesis.

Many thanks to all my lab mates; especially Gareth, Kyle and Iric for their patience and expertise. They made the lab a productive and friendly place to be.

I am also grateful for the financial support of the National Research Foundation for the final year of my PhD, which made my attendance at the 17th International Symposium on the Biology of Actinomycetes Conference possible. The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

List of Abbreviations

aa	-	amino acid
ANI	-	Average Nucleotide Identity
ATCC	-	American Type Culture Collection
BLAST	-	basic local alignment search tool
bp	-	base pair
CFU	-	colony forming unit
cm	-	centimeter
CZ	-	Czapek solution
DDH	-	DNA-DNA hybridization
DNA	-	deoxyribonucleic acid
dNTP	-	deoxyribonucleotide triphosphate
DSMZ	-	Deutsche Sammlung von Mikroorganismen und Zellkulturen
e.g.	-	*exempli gratia* - for the sake of example
et al	-	*et alia* - and others
g	-	gram
G+C	-	guanine + cytosine
h	-	hour
i.e.	-	*id est* – that is
IPTG	-	Isopropyl β-D-1-thiogalactopyranoside
ISP2	-	International Streptomyces Project 2
KCTC	-	Korean Collection for Type Cultures
km2	-	kilometers squared
KSM	-	Kribbella selective medium
LL-DAP	-	LL-diaminopimelic acid
m	-	meter

Mb	-	megabase
min	-	minute
mg	-	milligrams
ml	-	milliliter
mm	-	millimeter
mM	-	millimolar
MLSA	-	Multilocus sequence analysis
ng	-	nanogram
nm	-	nanometer
nt	-	nucleotide
OD	-	Optical density
PCR	-	Polymerase chain reaction
rRNA	-	ribosomal ribonucleic acid
s	-	second
SM1	-	Selective medium 1
sp. nov.	-	*species nova* - new species
ssDNA	-	single stranded deoxyribonucleic acid
subsp.	-	subspecies
U	-	units
w/v	-	weight for volume
X-gal	-	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
7H9	-	Middlebrook 7H9
°C	-	degrees Celsius
µg	-	microgram
µl	-	microliter
µM	-	micromolar

Abstract

Culture-dependent methods were used to isolate *Kribbella* strains from a soil sample, as well as leaf and root samples from 5 indigenous South African plants. To enhance the chances of isolating this rare actinobacterium, the development and use of a *Kribbella*-selective medium was employed. The leaf and root samples were surface sterilized before processing to isolate endophytes. All plates were incubated at 30°C (with a duplicate set at room temperature) for 12 weeks and any colonies which had a *Kribbella*-like morphology were sub-cultured for further study. The soil sample produced 6 strains identified as *Kribbella*. Further testing and sequence analysis using the 16S rRNA and *gyrB* genes indicated that all 6 strains were identical to the type strain of *Kribbella solani*. The plant samples did not produce any colonies identified as *Kribbella*. The *Kribbella* selective medium, although not successful in isolating *Kribbella* strains in this study, was able to enhance the selection of *Kribbella* from a soil sample during an Honours project. Thus, it is suggested that the *Kribbella* selective medium be evaluated further by including it in future *Kribbella* isolation projects.

A metagenomic method was then developed to indicate the presence of *Kribbella* DNA in the soil and the five plant samples previously screened using plate culture techniques. This culture-independent approach can be applied to a sample to ensure the presence of the bacteria before selective isolation on agar plates is attempted. A *Kribbella*-specific 16S rRNA gene PCR primer (Krb977F) was designed and the primer combination Krb977F+R5 (where R5 is a universal 16S rRNA gene primer) was used to amplify *Kribbella* 16S rRNA gene sequences directly from DNA extracted from soil and plant samples. Each amplified PCR product was cloned into a plasmid and transformed into *Escherichia coli* cells. The *E. coli* colonies were screened by PCR to ensure the correct sized insert was present. Transformants containing inserts were grown in broth culture and the plasmids were extracted and sequenced. The edited sequences were analyzed by BLAST against the GenBank and Ez-Taxon-e databases to determine the bacterium in the environmental sample from which the 16S rRNA gene sequence had been cloned.

Kribbella 16S rRNA gene sequences were successfully amplified from two soil samples and each of five leaf and five root samples. The Krb977F primer proved to be highly specific for *Kribbella* sequences, as only nine sequences belonging to cultured type strains of non-*Kribbella* genera were amplified. Of a total of 330 metagenomic 16S rRNA gene sequences, 181 (54.8%) were identified as belonging to *Kribbella*, 9 (2.7%) were identified as non-*Kribbella* cultured strains and 140 (42.4%) were most closely related to uncultured strains. The data also showed that *Kribbella* sequences were detected in both leaf and root samples of all the sampled indigenous plants (*Agathosma ovata* 'Glentana', *Leucospermum conocarpodendron*, *Orphium frutescens, Podylaria sericea* and *Protea aurea* subspecies *aurea*).

A multilocus sequence analysis (MLSA) method was developed to refine the phylogenetic analysis of the genus *Kribbella*. Sequences were obtained for the 16S rRNA, *gyrB, rpoB, recA, relA* and *atpD* genes for 20 type strains of the genus plus seven non-type strains in addition to four laboratory strains that were determined to represent new species. A five-gene concatenated sequence of 4099 nt was used to examine the phylogenetic relationships between the species of the genus *Kribbella*. Using the concatenated sequence of the *gyrB, rpoB, recA, relA* and *atpD* genes, most *Kribbella* type strains can be distinguished by a genetic distance of >0.04. For the four *Kribbella* species for which multiple strains are available (*Kribbella aluminosa*, *Kribbella catacumbae*, *Kribbella sancticallisti* and *K.solani*), the type strains share sequence similarities of 100% with the non-type strains of the same species for all genes except the 16S rRNA gene. By using MLSA with the *gyrB*, *rpoB, recA, relA* and *atpD* genes, an improved tree structure was obtained with increased bootstrap support, longer branch lengths and an overall higher resolution of the species in the genus. It is proposed that concatenated *gyrB-rpoB-recA-relA-atpD* gene sequences be used for examining the phylogenetic relationships within the genus *Kribbella* and for determining the closest phylogenetic relatives to be used for taxonomic comparison against new *Kribbella* isolates. Furthermore, the 0.04 concatenated-gene genetic distance threshold can be used as a species delineation tool in the genus *Kribbella*.

Chapter 1:
Introduction

Chapter 1: Introduction

Chapter 1: Introduction

1.1 Taxonomy

Taxonomy is a fundamental discipline underpinning the science of microbiology [Sutcliffe *et al*., 2012]. It consists of three key elements, namely, classification, nomenclature and identification, which provide a stable framework to sustain advances in biotechnology, clinical diagnostics, ecology and evolution [Prescott *et al*., 2002; Sutcliffe *et al*., 2012; Tindall *et al*., 2010]. Staley [2010] suggested that there is a fourth goal to taxonomy, namely, comprehending microbial diversity. Classification allows for the arrangement of organisms into specific groups (taxa) on the basis of shared characteristics [Kämpfer and Glaeser, 2012; Prescott *et al*., 2002]. Nomenclature assigns names to the taxonomic groups, which must be in agreement with published rules. Identification is the process of determining if an unknown isolate belongs to a recognized taxon [Kämpfer and Glaeser, 2012; Prescott *et al*., 2002]. Taxonomy is essential for the organization of huge amounts of knowledge about organisms [Prescott *et al*., 2002]. It places organisms into meaningful groups in which predictions about the organism can be made and gives organisms precise names to allow for efficient communication and identification [Prescott *et al*., 2002].

The full hierarchal taxonomic ranks are listed in Table 1.1, using *Kribbella flavida* as an example. The basic taxonomic group is the rank of species. In higher organisms, a species is defined as a group of interbreeding or potentially interbreeding natural populations that are reproductively isolated from other groups [Prescott *et al*., 2002]. The term species is defined differently in bacteriology in which a species is considered to be a collection of strains that have a similar G+C% composition and 70% or greater DNA relatedness as judged by DNA-DNA hybridization (DDH) analysis [Prescott *et al*., 2002].

Table 1.1 Full taxonomic classification of type species *Kribbella flavida*

Rank	
Domain	*Bacteria*
Phylum	*Actinobacteria*
Class	*Actinobacteria*
Subclass	*Actinobacteridae*
Order	*Actinomycetales*
Suborder	*Propionibacterineae*
Family	*Nocardioidaceae*
Genera	*Kribbella*
Species	*flavida*

DDH has become the gold standard for species delineation as defined by Wayne *et al*. [1987], where a species includes strains with 70% or greater DNA relatedness and with a difference of 5°C or less in the melting temperature of the DNAs of the members of the species. Furthermore, strains of a species must have a distinct phenotype from closely related species [Staley, 2006; Wayne *et al.,* 1987]. Prokaryotic species are characterized by a variety of phenotypic differences and by defined genotypic threshold values such as DDH values or 16S rRNA gene similarity values [Staley, 2006].

1.1.1 History of Taxonomy

Prior to the 1960's, early classifications of bacteria depended entirely on the phenotypic properties expressed by the organism with great emphasis on metabolic and physiological features [Richter and Róssello-Móra, 2009; Staley, 2006; Staley, 2009]. The phenotypic attributes included growth temperature, pH range for growth, carbon sources utilized, habitat, Gram stain, size and cell shape [Staley, 2009]. It was these attributes and traits that had formed the basis for early classifications such as those found in Bergey's Manual of Systematic Bacteriology [Staley, 2006]. Since the late 1960's there have been many technical innovations [Staley, 2009]. The first innovation being the determination of mol% G+C content and second, the use of DNA-DNA hybridization, which would later become the basis for species delineation [Busse *et al*., 1996; Richter and Róssello-Móra, 2009]. The next important innovation was the work pioneered by Carl Woese and colleagues, which used 16S rRNA gene sequences and introduced phylogenetic analyses to examine evolution through the sequences of proteins and DNA [Busse *et al*., 1996].

The current modern practice of taxonomy is referred to as a polyphasic approach. This approach uses multiple aspects of an organism, which include phenotypic and chemotaxonomic traits as well as genotypic and phylogenetic information [Chun and Rainey, 2014; Coenye *et al*., 2005]. The combination of traditional phenotypic and chemotaxonomic tests with genetic information allows for a more comprehensive classification of an organism [Sutcliffe *et al*., 2012].

1.1.2 DNA-DNA Hybridization

DDH is an experimental method to measure the degree of relatedness between two different genomes by hybridizing the genomes together [Chun and Rainey, 2014]. The principle of DDH measures the amount of binding between complementary regions of single stranded DNA of two different origins [Ludwig, 2007]. One method involves incubating bound single stranded DNA (ssDNA) from one strain with radioactively labelled ssDNA from another strain, allowing hybridization to occur and measuring the amount of newly bound radioactive DNA [Prescott *et al*., 2002]. The amount of label present reflects the amount of hybridization and thus the similarity of the sequences. The degree of similarity is expressed as the % of labelled DNA retained on the filter compared to the % of homologous DNA bound under the same conditions [Prescott *et al*., 2002].

Although DDH has been the gold standard with regards to delineating species there are some drawbacks. DDH is a complex, time consuming and labour intensive technique and the various methods produce different results [Goris *et al*., 2007; Richter and Róssello-Móra, 2009]. However, the main drawback is that, due to the comparative nature of the method, a cumulative database of DDH results is not possible and this has encouraged taxonomists to search for alternative methods [Goris *et al*., 2007; Richter and Róssello-Móra, 2009].

1.1.3 Use of the 16S rRNA gene

Another important innovation was the use of rRNA sequences to classify bacteria by evolutionary means that was pioneered by Woese in 1977 [Ludwig, 2007]. The use of

16S rRNA gene sequences allowed the creation of a hierarchical taxonomic system based on one practical, universally distributed trait and the use of 16S rRNA gene sequences now forms the backbone of prokaryote taxonomy [Stackebrandt *et al*., 1997; Gevers *et al*., 2005; Zhi *et al*., 2009; Kämpfer, 2012; Kämpfer and Glaeser, 2012]. The sequencing of 16S rRNA genes is now a widely accepted primary technique to identify bacteria to the genus level and is generally the first test used to identify a novel isolate [Staley, 2006]. More recently, it has been proposed that a pairwise 16S rRNA gene sequence similarity of < 98.7% 16S rRNA gene sequence similarity can be considered to coincide with the 70% DDH cut off for distinguishing between genomic species [Stackebrandt and Ebers, 2006; Meier-Kolthoff *et al*., 2013]. Thus, DDH is not required to differentiate between two strains if their 16S rRNA gene sequence similarity is< 98.7% [Sutcliffe *et al*., 2012]. For over 20 years bacterial taxonomists have primarily relied on 16S rRNA gene sequence analyses for identification and classification [Rajendhran and Gunasekaran, 2011]. However, there are some drawbacks to the use of 16S rRNA gene sequences for phylogenetic analysis, as this gene has a slow evolutionary rate, it spans a very small part of the genome, it does not encode a protein (and therefore any insertions or deletions in the sequence may distort actual relationships), and it is too highly conserved to distinguish between closely related species [Santos and Ochman, 2004; Staley, 2009].

1.1.4 Multilocus sequence analysis (MLSA)

Due to the shortcomings of the 16S rRNA gene in differentiating between closely related species and the fact that most bacterial phylogenetic analyses are based solely on this gene, alternatives have been sought out. The use of MLSA has been shown to improve upon the phylogenetic analysis of many genera of actinobacteria, including the genus *Kribbella* [Curtis and Meyers, 2012].

MLSA is based on multilocus sequence typing (MLST), which is based on sequences of multiple genes and was first introduced by Maiden and colleagues [Maiden *et al*., 1998] to investigate population structures of pathogenic microorganisms [Kämpfer and Glaeser, 2012]. MLSA uses the sequence data behind MLST for phylogenetic analyses and can be used for bacterial identification as well as classification [Kämpfer and Glaeser, 2012].

MLSA utilizes a concatenated sequence of several in-frame protein coding gene fragments on which to base the phylogenetic analysis [Richter and Róssello-Móra, 2009]. The genes selected for MLSA are usually ubiquitous, universally conserved housekeeping genes, which are present in a single copy, are over 900 bp in length and are well separated on the genome [Adékambi*et al*., 2011; Santos and Ochman, 2004; Zeigler, 2003]. By using multiple genes, the discriminatory power and robustness of the trees increases [Guo *et al*., 2008]. MLSA has been shown to determine a more accurate evolutionary history and is a useful tool for distinguishing between species [Carro *et al*., 2012; Dalmasso *et al*., 2011; Rajendhran and Gunasekaran 2011; Staley 2006]. For the genus *Kribbella*, the MLSA approach was able to resolve the species *Kribbella solani* and *Kribbella hippodromi*, which could not be distinguished by 16S rRNA gene analysis. In addition, the MLSA also increased the confidence in the topology of the tree, as well as the resolution [Curtis and Meyers, 2012].

However, there are some downsides to MLSA. Firstly, there is currently no consensus list of the genes to be used. Secondly, one must design primers to amplify each gene (although the rapid increase in the number of bacterial genome sequences offers the hope that some of this sequence-acquisition work will be reduced). Thirdly, a bias may be introduced into the analysis, as the whole genome is not utilized, through the gene selection process of what specific genes are used and how many are chosen for the analysis [Richter and Róssello-Móra, 2009]. As an alternative to DDH for species delineation, MLSA may have to be applied and assessed for each genus. Despite these issues, MLSA can be made to be portable by depositing the sequences in publicly accessible databases, it has shown higher resolution than the 16S rRNA gene and represents the genome more accurately by using multiple genes [Ludwig, 2007]. As DNA sequencing is becoming less expensive and more accessible, MLSA provides a good alternative to DDH, as well as more refined phylogenetic analyses than using the 16S rRNA gene alone.

1.1.5 Future prospects of Taxonomy

Although 16S rRNA gene analyses and DDH continue to be the main molecular criteria for species delineation, it is expected that much more taxonomic information can be

extracted from complete genome sequences [Coenye *et al*., 2005]. Genome sequencing can greatly assist standardizing taxonomy because it offers higher accuracy compared to the traditional morphological and biochemical observations and will enable a better understanding of the role of expression in determining phenotypic patterns [Konstantinidis and Tiedje, 2007]. The rapid advances in molecular biology will likely lead to a sequence based approach, in which, full genome comparisons between strains are made possible [Gevers *et al.,* 2005]. Already, *Antonie van Leewenhoek Journal of Microbiology* encourages authors to provide the genome sequence of the type strain of a proposed new species [Antonie van Leeuwenhoek, 2014].

In the era of genomics, where genetic information can be accessed in public databases, DDH has become an obsolete approach that urgently needs substitution [Richter and Róssello-Móra, 2009]. Currently, the use of Average Nucleotide Identity (ANI) seems the best alternative to easily replace DDH as the gold standard [Chun and Rainey, 2014; Richter and Róssello-Móra, 2009]. ANI is a genome sequence derived relatedness index that produces the mean of a set of similarity values from orthologous regions of two genomes [Chun and Rainey, 2014]. It has been found that ANI values correlate well with DDH values and that an ANI value of 95-96% is equivalent to the 70% DDH cut-off threshold [Chun and Rainey, 2014; Richter and Róssello-Móra, 2009].

Moving forward, whole genome sequencing will provide important insights into physiology and biotechnological potential of bacterial strains [Sutcliffe *et al*., 2012]. It will also allow for the delineation of species to be assessed more comprehensively and illuminate the impact of lateral genome transfer on prokaryotic phylogeny [Sutcliffe *et al*., 2012]. Modern taxonomy was greatly advanced by the introduction of 16S rRNA gene analyses and genomics will have an equal impact [Chun and Rainey, 2014]. The comparison of whole genome sequences will allow great insight into the ecology, physiology and taxonomy of microorganisms and with these advances, should be able to provide more complete and objective approaches for classification and identification [Chun and Rainey, 2014; Sutcliffe *et al*., 2012].

1.2 *Actinobacteria*

In terms of number and variety of identified species, the actinobacteria, belonging to the phylum *Actinobacteria*, represents one of the largest taxonomic units among the domain Bacteria [Ventura *et al*., 2007]. The *Actinobacteria* are Gram positive and are known for having a high G+C content in their DNA, ranging from 51%-70% [Adékambi *et al*., 2011]. They have a wide distribution and are found in terrestrial, marine and freshwater aquatic ecosystems [Ventura *et al*., 2007]. In soil, the actinobacteria play a crucial role in biodegradation (e.g. degrading lignin) and bioremediation (e.g. degrading recalcitrant aromatic compounds) [Guo *et al*., 2008; Ventura *et al*., 2007]. They exhibit diverse phenotypic characteristics and morphologies, from cocci and rods to highly-branched filamentous forms. The filamentous actinobacteria are commonly referred to as actinomycetes. Some of the genera of filamentous actinomycetes exhibit fragmentation in which the hyphae break up into coccoid or rod-shaped elements in liquid and/or agar cultures (e.g. *Amycolatopsis* and *Kribbella*). Actinobacteria also exhibit diverse physiological and metabolic properties, such as the production of extra-cellular enzymes and the formation of a wide variety of secondary metabolites [Ventura *et al*., 2007].

Currently the Phylum and Class of *Actinobacteria* is made up of 5 subclasses, 10 orders 17 suborders and 56 families Table 1.2) and a total of 341 genera [Parte, 2014]. The Order *Actinomycetales* is the largest order and is the order to which many important actinobacteria belong. The *Actinobacteria* contain not only pathogenic bacteria (e.g. *Mycobacterium tuberculosis*), but also biotechnologically and pharmaceutically important strains (e.g. antibiotic producing strains of *Streptomyces*) [Kurtböke, 2012]. Actinobacteria are the most economically significant prokaryotes, producing about 45% of all the bioactive compounds produced by microorganisms [Kurtböke, 2012; Qin *et al*., 2009]. Amazingly, one genus *Streptomyces,* is responsible for producing 80% of all microbial natural products [Kurtböke, 2012]. The genus *Streptomyces* is very commonly found in soil environments and remains unsurpassed in the ability to produce secondary metabolites [Kurtböke, 2012]. However, with antibiotic resistance increasing, there is a strong need to look for new and different natural compounds. One strategy to avoid rediscovering known compounds is to search for new actinobacterial species, as screening actinobacterial biodiversity should provide access to novel secondary metabolites.

Table 1.2. The Classification of the Phylum and Class *Actinobacteria* down to Family level, adapted from [Parte, 2014].

Subclass	Order	Suborder	Family
Acidimicrobidae	*Acidimicrobiales*	*Acidimicrobineae*	*Acidimicrobiaceae*
			Iamiaceae
Actinobacteridae	*Actinomycetales*	*Actinomycineae*	*Actinomycetaceae*
		Actinopolysporineae	*Actinopolysporaceae*
		Catenulisporineae	*Actinospicaceae*
			Catenulisporaceae
		Corynebacterineae	*Corynebacteriaceae*
			Dietziaceae
			Mycobacteriaceae
			Nocardiaceae
			Segniliparaceae
			Tsukamurellaceae
		Frankineae	*Acidothermaceae*
			Cryptosporangiaceae
			Geodermatophilaceae
			Frankiaceae
			Motilibacteraceae
			Nakamurellaceae
			Sporichthyaceae
		Glycomycineae	*Glycomycetaceae*
		Jiangellineae	*Jiangellaceae*
		Kineosporiineae	*Kineosporiaceae*
		Micrococcineae	*Beutenbergiaceae*
			Bogoriellaceae
			Brevibacteriaceae
			Cellulomonadaceae
			Demequinaceae
			Dermabacteraceae
			Dermacoccaceae
			Dermatophilaceae
			Intrasporangiaceae
			Jonesiaceae
			Microbacteriaceae
			Micrococcaceae
			Promicromonosporaceae
			Rarobacteraceae
			Ruaniaceae
			Sanguibacteraceae
		Micromonosporineae	*Micromonosporaceae*
		Propionibacterineae	*Nocardioidaceae*
			Propionibacteriaceae
		Pseudonocardineae	
			Pseudonocardiaceae
		Streptomycineae	*Streptomycetaceae*
		Streptosporangineae	*Nocardiopsaceae*
			Streptosporangiaceae
			Thermomonosporaceae
	Bifidobacteriales		*Bifidobacteriaceae*
Coriobacteridae	*Coriobacteriales*	*Coriobacterineae"*	*Coriobacteriaceae*
Nitriliruptoridae			
	Euzebyales		*Euzebyaceae*
	Nitriliruptorales		*Nitriliruptoraceae*
Rubrobacteridae			
	Gaiellales		*Gaiellaceae*
	Rubrobacterales	*Rubrobacterineae*	*Rubrobacteraceae*
	Solirubrobacterales		*Conexibacteraceae*
			Patulibacteraceae
			Solirubrobacteraceae
	Thermoleophilales		*Thermoleophilaceae*

1.2.1 The genus Kribbella

The genus *Kribbella* is considered one of the rarer genera within the actinomycetes. It is a relatively young genus, having been established in 1999 by Park *et al.* [1999]. The genus currently comprises of 20 species with validly published names and has been steadily growing [Parte, 2014]. *Kribbella* is the second largest genus in the family *Nocardioidaceae* (sub-order *Propionibacterineae*) after the type genus *Nocardioides* (73 species with validly published names) [Park *et al*., 1999; Parte, 2014].

The vegetative mycelium of *Kribbella* species comprises extensively branched hyphae that may fragment into small rod-shaped elements in liquid cultures [Park *et al*., 1999]. *Kribbella* colonies are cream to beige in colour with lichenous wrinkled shapes with irregular edges (Figure 1.1). The predominant menaquinone is MK-9(H_4) and the genomic DNA G+C content ranges from 66-77 mol % [Park *et al*., 1999; Sohn *et al*., 2003; Everest *et al*., 2013]. The diamino acid in the peptidoglycan is LL-diaminopimelic acid and the major cellular fatty acid is anteiso-$C_{15:0}$ [Park *et al*., 1999]. The major polar lipids are phosphatidylcholine and phosphatidylglycerol. There are no diagnostic sugars present in the whole cell sugar pattern [Everest *et al*., 2013].

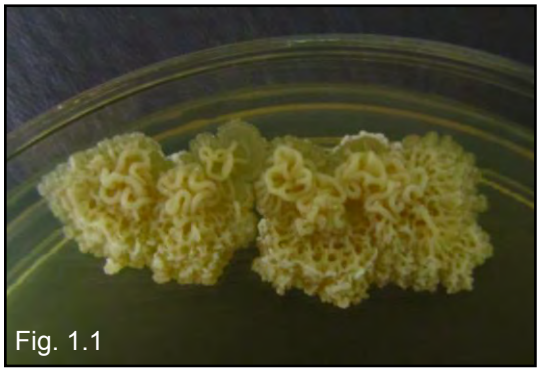


Figure. 1.1. Typical wrinkled morphology of *Kribbella* colonies on an agar plate. The first signs of sporulation are visible as a white rim on a colony at the top of the picture.

Kribbella species have been isolated from a variety of sources, with most species isolated from terrestrial soil. *K. hippodromi*, *Kribbella swartbergensis* and *Kribbella karoonensis* were all isolated from South African soil [Everest and Meyers, 2008; Kirby

et al., 2006]. Other interesting isolation sources include the walls of a catacomb (*Kribbella catacumbae, Kribbella sancticallisti, Kribbella albertanoniae* and *Kribbella italica*) [Urzì *et al*., 2008; Everest *et al*., 2013; Everest *et al*., 2015], and from rock surfaces in an alum slate mine (*Kribbella aluminosa*) [Carlsohn *et al*., 2007]. *Kribbella* species have also been isolated from an overlooked and understudied environmental niche, i.e. living plants. These endophytic species include *Kribbella lupini* isolated from the roots of the plant *Lupinus angustifolius* [Trujillo *et al.,* 2006], *Kribbella endophytica* from the leaves of the tree *Pittosporum angustifolium* [Kaewkla and Franco, 2013] and '*Kribbella podocarpi'*, which was isolated from a South African Yellowwood tree, *Podocarpus latifolius*, by a previous Masters student in the laboratory [Curtis *et al*., manuscript under revision].

1.3 Endophytes and fynbos plants

1.3.1 Endophytes

Soil has been the predominant and conventional source of actinomycete strains and contains great diversity, as well as a large density of actinomycetes. It has been found, however, that living plant material is an under-explored source of actinomycete biodiversity [Janso and Carter, 2010; Qin *et al*., 2011]. The term given to a microorganism that resides in the tissue of living plants is endophyte [Hasegawa et *al*., 2006; Qin *et al*., 2011]. Using a stricter definition, a true endophyte must be microscopically observed in the plant tissue in addition to having been isolated from surface sterilized plant material and have the ability to re-colonize plant tissue [Hasegawa *et al*., 2006; Rosenblueth and Martínez-Romero, 2006]. The observation of the bacteria in the plant tissue is not always fulfilled and the term putative endophyte has been proposed, yet putative endophytes compose most of the endophytes reported to date. In this thesis the term endophyte will be used for both true and putative endophytes [Hasegawa *et al*., 2006; Rosenblueth and Martínez-Romero, 2006].

Endophytes have proven to be ubiquitous and it has been demonstrated that some endophytes actually improve and promote growth in the host plants, as well as reduce disease symptoms caused by pathogens or environmental stress [Hasegawa, *et al.,* 2006; Qin *et al*., 2009]. It has long been established that actinobacteria are closely

associated with living plants, in addition to other conventional environments such as soil and pond sediment, with the first actinobacterial endophyte (*Frankia* sp.) isolated from a non-legume root nodule in 1886 [Hasegawa *et al* , 2006].

Almost all vascular plants examined to date have been found to harbour several strains of endophytic actinomycetes, which shows that plant-associated actinomycetes are quite common [Hasegawa *et al*., 2006]. The bacteria living in the plant receive protection and nutrition, while supplying a variety of bioactive metabolites to enhance the fitness of the host plant [Hasegawa *et al*., 2006]. The promotion of plant growth by the endophyte may be due to nitrogen fixation, the production of phytohormones, or the biocontrol of plant pathogens by the production of compounds such as antibiotics and antifungals [Hasegawa *et al*., 2006; Rosenblueth and Martínez-Romero, 2006]. Endophytic actinomycetes have the potential to produce a wide range of natural compounds from antibiotics, enzymes and plant growth promoters, which may be exploited for agricultural and pharmaceutical uses [Hasegawa, *et al.,* 2006]. One example is the discovery of the novel compound, coronamycin, which was found to be produced by an endophytic *Streptomyces* species [Ezra *et al*., 2004]. The *Streptomyces* strain was isolated from the stem of a vine (genus *Monstera*) from the Amazon jungle in Peru [Ezra *et al*., 2004]. The strain was found to have antifungal activity against plant pathogens, but had the best bioactivity against *Plasmodium falciparum*, the parasite responsible for malaria [Ezra *et al*., 2004].

The discovery of coronamycin from the Peruvian Amazon also serves as an example that, in ecosystems with a high general biodiversity, it is likely to also find a great diversity of endophytes [Strobel *et al*., 2004]. Many investigations have targeted plants from tropical areas, since tropical rainforests are very biologically diverse and thus provide the greatest possible chance of finding novel actinobacteria and novel compounds [Janso and Carter, 2010; Qin *et al*., 2009]. One other example of success was the discovery of a novel antibiotic named spoxazomicin produced by *Streptosporangium oxazolinicum* K07-0450^{T}, which was isolated from an orchid in the subtropical Okinawa region of Japan [Brader *et al*., 2014; Inahashi *et al*., 2011].

1.3.2 Fynbos plants and the Cape Floral Kingdom

Guided by the idea that plants growing in areas of great biodiversity have the prospect of harbouring endophytes with great biodiversity, the Cape Floral Kingdom proves an ideal location to study [Ezra *et al*., 2004]. The Cape Floral Kingdom or Cape Floristic Region (outlined in Figure 1.2 and in detail in Figure 2.1) is, for its size (90,000km^2), the most diverse plant kingdom [Goldblatt and Manning, 2002; Taylor *et al.,* 2001]. It has the world's highest representation of endemic plants with an estimated 9030 vascular plant species (68.7% endemic) [Goldblatt and Manning, 2002; Marean, 2010; Taylor *et al.,* 2001]. The Cape Floristic Region also comprises 44% of the estimated total of 20,500 species that occur in all of southern Africa and this level of species richness compares to the richness found in the wet tropics rather than the temperate zone in which it occurs [Goldblatt and Manning, 2002].

Of the many vegetation regions within the Cape Floral Kingdom (Figure 2.1), the dominate vegetation type is the fynbos, representing an estimated 70% of the vascular plants in the Cape Floral Kingdom [Brown, 1993; Rebelo and Siegfried, 1990]. Fynbos manifests as shrubland or heathland, occurs in sandy soils with an annual rainfall of 250mm to 700mm and is regenerated by fire [Rebelo and Siegfried, 1990]. It is from this vegetation group that the five plants (*Agathosma ovata* 'Glentana', *Orphium frutescens, Leucospermum conocarpodendron, Podylaria sericea* and *Protea aurea* subspecies *aurea*, pictured in Figure 2.2) were chosen for the study.

The Cape Floral Kingdom is an ideal location from which to obtain a diverse set of plants in which to look for novel endophytic strains of actinomycetes. As the frequency of discovering novel compounds is decreasing, new microbial habitats need to be examined and thus endophytic actinomycetes from a highly biodiverse ecosystem may be a promising source of new natural products [Janso and Carter, 2010; Qin *et al*., 2011; Strobel *et al.,* 2004].

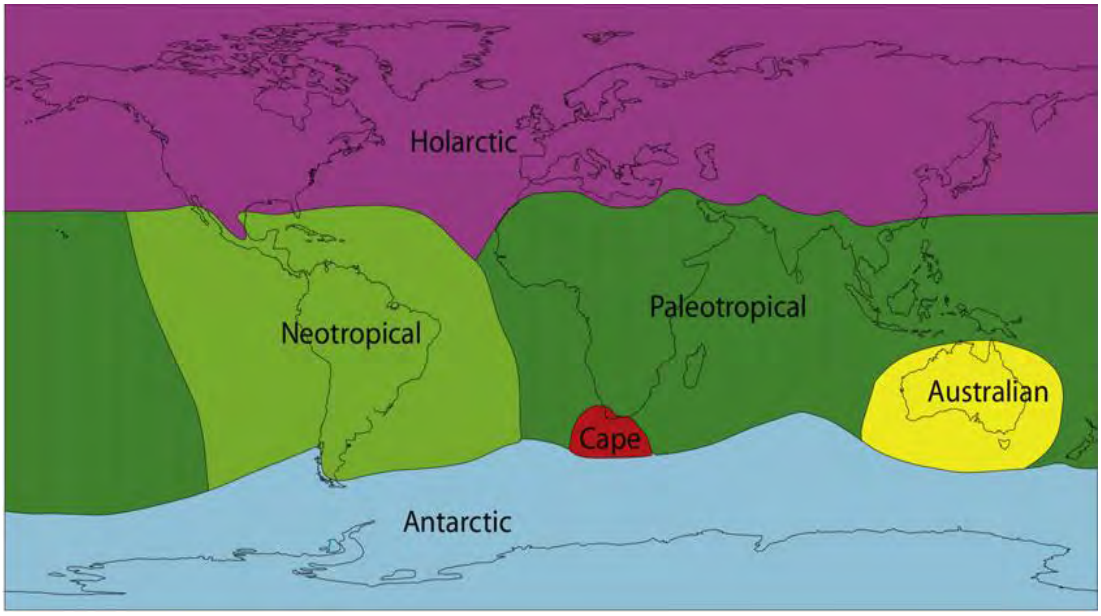


Figure 1.2 Global map of the Earth's 6 Floral Kingdoms. The smallest of the Kingdoms, the Cape Floristic Region, is indicated in red [Marean, 2010].

1.4 Metagenomics

The majority of microbes in nature have not been studied or cultured under laboratory conditions and it is estimated that culturable microbes account for only 1% or less of the bacterial diversity in most environmental samples [Riesenfeld *et al.,* 2004; Rondon *et al.,* 2000; Schmeisser *et al.,* 2007]. The hidden diversity of the uncultured 99% represents an intriguing resource for the development of novel enzymes, antibiotics and other secondary metabolites for use in biotechnology. Metagenomic analysis allows this hidden diversity to be assessed [Riesenfeld *et al.,* 2004; Schmeisser *et al.,* 2007].

Metagenomics is the tool by which the total genetic pool of all the microbes in a particular environment may be investigated in a culture-independent manner [Culligan *et al*., 2014; Schmeisser *et al*., 2007]. The term is derived from the statistical concept of meta-analysis and genomics and involves cloning large fragments of DNA isolated directly from microbes in natural environments [Rondon *et al*., 2000; Schloss and Handelsman, 2003; Schmeisser *et al.,* 2007]. After extracting DNA directly from an environmental sample, it is cloned using an appropriate vector and inserted into a culturable bacterium [Daniel, 2004]. This clone library can now be screened for

functional genes (e.g. using antibiotic activity screening assays) or a sequence based approach may be used where PCR primers are designed to target specific genes as outlined in Figure 1.3 [Daniel, 2004].

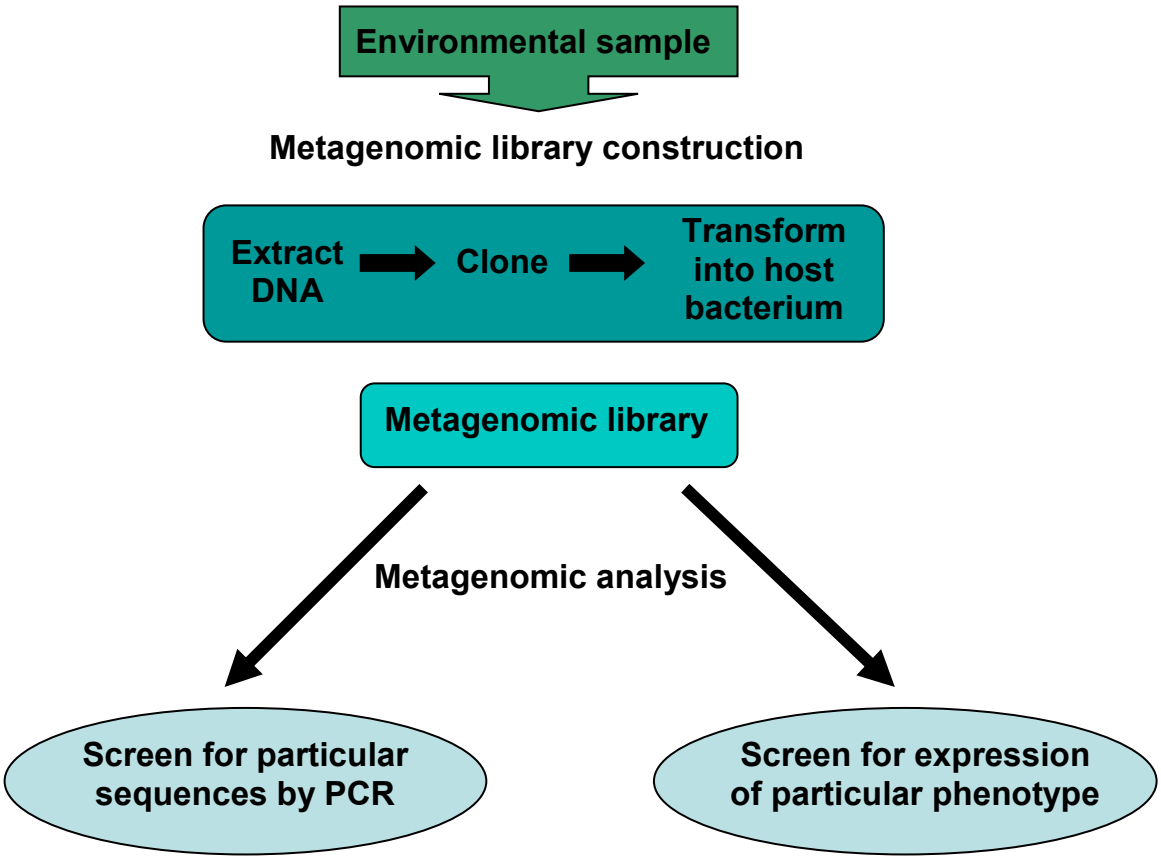


Figure. 1.3 Outlined process of metagenomic analysis, illustrating the sequence-based or functional expression approach. Adapted from Riesenfeld *et al*. [2004].

Metagenomics has been used successfully as a tool on a variety of levels, from the study of single genes, pathways, organisms and entire communities [Riesenfeld *et al*., 2004]. Metagenomics has been invaluable in the study of the human gut microbiome and has lead to greater understanding of changes in the gut's microbial communities under various circumstances [Graessler *et al.,* 2013; Lepage *et al*., 2013; Levy and Borenstein, 2014; Raes, 2014; Turnbaugh *et al.,* 2009].

There have been many studies where novel antibiotics were detected in metagenomic libraries including novel isocyanide-containing antibiotics from a soil metagenomic library [Schmeisser *et al*., 2007]. In this study, metagenomic analysis using a sequence based approach was used to determine the presence of a single actinobacterial genus living in soil and in plant tissues.

1.5 Aims

It was the aim of this study to investigate 5 indigenous fynbos plants from the biodiverse Cape Floral Kingdom for the presence of endophytic *Kribbella* strains living in the leaf and root tissues. In addition, a soil sample was investigated for the presence of *Kribbella* strains using plate cultivation techniques. A selective medium was designed to favour the growth of *Kribbella* strains in order to further enhance the search for *Kribbella*. In addition to the culture-dependant methods, a metagenomic screening method was developed to detect the presence of *Kribbella* DNA in the environmental samples with the use of a *Kribbella*-specific 16S rRNA gene PCR primer. *Kribbella* is an ideal candidate to search for since it has been isolated from a wide variety of habitats and, although not a well-known producer of antibiotics, novel strains could prove to be producers of novel secondary metabolites.

A further aim of this study was to refine the understanding of the phylogenetic relationships of the species in the genus *Kribbella* and to provide additional information towards a DNA sequence-based replacement for DDH in defining new *Kribbella* species. The use of one gene (*gyrB)* by Kirby *et al.*, [2010] was shown to improve upon the standard 16S rRNA gene phylogenetic analysis of the genus and thus a multilocus sequence analysis for the genus was developed using multiple housekeeping genes. By using MLSA with the *gyrB*, *rpoB, recA, relA* and *atpD* genes an improved tree structure was obtained with increased bootstrap support and branch lengths and an overall higher resolution of the species in the genus. This study also proposes the use of a *Kribbella* species delineation tool utilizing a genetic distance threshold.

By investigating indigenous plants for the presence of *Kribbella* strains by using both culture-dependent and culture-independent methods, further insights into the ecology of the genus can be assessed and the chance of the discovery of novel natural products increases. Furthermore, with the tools developed through the MLSA and genetic distance threshold, the phylogenetic analysis of the genus has been refined and any new *Kribbella* isolates may be quickly assessed to determine whether they are likely to represent novel species. This not only improves upon the taxonomy of the genus, but enables new species to be characterized with a more robust phylogenetic analysis and potentially without DDH.

1.6 References

Adékambi, T., Butler, R.W., Hanrahan, F., Delcher, A.L., Drancourt, M. and Shinnick, T.M. (2011). Core gene set as the basis of Multilocus Sequence Analysis of the subclass Actinobacteridae. *PLoS ONE*. **6**, e14792.

Antonie van Leeuwenhoek (2014). Instructions for Authors. (http://www.springer.com/life+sciences/microbiology/journal/10482). Accessed September 2014.

Brader, G., Compant, S., Mitter, B., Trognitz, F. and Sessitsch, A. (2014). Metabolic potential of Endophytic bacteria. *Curr. Opin. Microbiol*. **27,** 30-37.

Brown, N.A.C. (1993). Promotion of germination of fynbos seeds by plant derived smoke. *New Phytol*. **123,** 575-583.

Busse, H.-J., Denner, E.B.M. and Lubitz, W. (1996). Classification and identification of bacteria: current approaches to an old problem. Overview of methods used in bacterial systematics. *J. Biotechnol.* **47**, 3-38.

Carlsohn, M.R., Groth, I., Spröer, C., Schütze, B., Saluz, H.-P., Munder, T., and Stackebrandt, E. (2007). *Kribbella aluminosa* sp. nov., isolated from a medieval alum slate mine. *Int. J. Syst. Evol. Microbiol.* **57**, 1943-1947.

Carro, L., Spröer, C., Alonso, P. and Trujillo, M.E. (2012). Diversity of *Micromonospora* strains isolated from nitrogen fixing nodules and rhizosphere of *Pisumsativum* analyzed by multilocus sequence analysis. *Syst. Appl. Microbiol*. **35**, 73-80.

Chun, J. and Rainey, F.A. (2014). Integrating genomics into the taxonomy and systematics of the *Bacteria* and *Archaea*. *Int. J. Syst. Evol. Microbiol.* **64**, 316-324.

Coenye, T., Gevers, D., Van de Peer, Y., Vandamme, P. And Swings, J. (2005). Towards a prokaryotic genomic taxonomy. *FEMS Microbiol.Rev*. **29**, 147-167.

Culligan, E.P., Sleator, R.D., Marchesi, J.R., and Hill, C. (2014). Metagenomics and novel gene discovery: Promise and potential for novel therapeutics. *Virulence*. **5**, 399-412.

Curtis, S.M., Norton, I., Everest, G.J. and Meyers, P.R. (2015). *Kribbella podocarpi* sp. nov., isolated from the leaves of a yellowwood tree (*Podocarpus latifolius*) and *Kribbella capetownensis* sp. nov.and*Kribbella speibonae* sp. nov., isolated from soil. *Antonie van Leeuwenhoek.* Manuscript under revision.

Curtis, S.M. and Meyers, P.R. (2012). Multilocus sequence analysis of the actinobacterial genus *Kribbella*. *Syst. Appl. Microbiol.* **35**, 441-446.

Dalmasso, M., Nicolas, P., Falentin, H., Valence, F., Tanskanen, J., Jatila, H., Salusjärvi,T. and Thierry, A. (2011). Multilocus sequence typing of *Propionibacterium freudenreichii*. *Int. J. Food. Microbiol.* **145**, 113-120.

Daniel, R. (2004). The soil metagenome- a rich resource for the discovery of novel natural products. *Curr. Opin. Biotech*. **15,** 199-204.

Everest, G.J., Curtis, S.M., De Leo, F., Urzì, C. and Meyers, P.R. (2015). Description of *Kribbella italica* sp. nov., isolated from a Roman catacomb. *Int. J. Syst. Evol. Microbiol*. **65**, 000-000 (*In Press*).

Everest, G.J., Curtis, S.M., De Leo, F., Urzì, C. and Meyers, P.R. (2013). *Kribbella albertanoniae* sp. nov., isolatated from a Roman catacomb, and emended description of the genus *Kribbella*. *Int. J. Syst. Evol. Microbiol*. **63**, 3591-3596.

Everest, G.J., and Meyers, P.R. (2008). *Kribbella hippodromi* sp. nov., isolated from soil from a racecourse in South Africa. *Int. J. Syst. Evol. Microbiol*. **58**, 443-446.

Ezra, D., Castillo, U.F., Strobel, G.A., Hess, W.M., Porter, H., Jensen, J.B., Condron, M.A.M., Teplow, D.B., Sears, J., Maranta, M., Hunter, M., Weber, B. and Yaver, D. (2004). Coronamycins, peptide antibiotics produced by a verticillate *Streptomyces* sp. (MSU-2110) endophytic on *Monstera* sp. *Microbiology.* **150**, 785-793.

Gevers, D., Cohan, F.M., Lawrence, J.G., Spratt, B.G., Coenye, T., Feil, E.J., Stackebrandt, E., Van de Peer, Y., Vandamme, P., Thompson, F.L. and Swings, J. (2005). Re-evaluating prokaryotic species. *Nat. Rev. Microbiol.* **3,** 733-739

Goldblatt, P. and Manning, J.C. (2002). Plant Diversity of the Cape Region of Southern Africa. *Annals of the Missouri Botanical Garden*. **89**, 281-302.

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

Graessler, J., Qin, Y., Zhong, H., Zhang, J., Licinio, J., Wong, M.-L., Xu, A., Chavakis, T., Bornstein, A.B., Erhart-Bornstein, M., Lamounier-Zepter, V., Lohmann, T., Wolf, T., and Bornstein S.R. (2013). Metagenomic sequencing of the human gut microbiome before and after bariatric surgery in obese patients with type 2 diabetes: correlation with inflammatory and metabolic parameters. *Pharmacogenomics*. **13**, 514-522.

Guo, Y., Zheng, W., Rong, X. and Huang, Y. (2008). A multilocus phylogeny of the *Streptomyces griseus* 16S rRNA gene clade: use of multilocus sequence analysis for streptomycete systematics. *Int. J. Syst. Evol. Microbiol .***58**, 149-159.

Hasegawa, S., Meguro, A., Shimizu, M., Nishimura, T. and Kuno, H. (2006). Endophytic Actinomycetes and Their Interactions with Host Plants. *Actinomycetologica.* **20,** 72-81.

Inahashi, Y., Matsumoto, A., Omura, S. and Takahashi, Y. (2011). *Streptosporangium oxazolinicum* sp. nov., a novel endophytic actinomycete producing new antitrypanosomal antibiotics, spoxazomicins. *J. Antibiot*. **64**, 297–302.

Janso, J.E., Carter, G.T. (2010). Biosynthetic Potential of Phylogenetically Unique Endophytic Actinomycetes from Tropical Plants. *Appl. Environ. Microbiol.* **76**, 4377-4386.

Kaewkla, O. and Franco, C.M.M. (2013). *Kribbella endophytica* sp. nov., an endophytic actinobacterium isolated from the surface-sterilized leaf of a native apricot tree. *Int. J. Syst. Evol. Microbiol .***63**, 1249-1253.

Kämpfer, P. (2012). Systematics of prokaryotes: the state of the art. *Antonie van Leeuwenhoek.* **101**, 3-11.

Kämpfer, P. and Glaeser, S.P. (2012). Prokaryotic taxonomy in the sequencing era- the polyphasic approach revisited. *Environ. Microbiol*. **14**, 291-317.

Kirby, B.M., Everest, G.J. and Meyers, P.R., (2010). Phylogenetic analysis of the genus *Kribbella* based on the *gyrB* gene: proposal of a *gyrB*-sequence threshold for species delineation in the genus *Kribbella*. *Antonie van Leeuwenhoek*. **97**, 131-142.

Kirby, B.M., Le Roes, M., Meyers, P.R. (2006). *Kribbella karoonensis* sp. nov. and *Kribbella swartbergensis* sp. nov., isolated from soil from the Western Cape, South Africa. *Int. J. Syst. Evol. Microbiol* .**56**, 1097-1101.

Konstantinidis, K.T. and Tiedje, J.M. (2007). Prokaryotic taxonomy and phylogeny in the genomic era: advancements and challenges ahead. *Curr. Opin. Microbiol*. **10**, 504-509.

Kurtböke, D. Í. (2012). Biodiscovery from rare actinomycetes: an eco-taxonomical perspective. *Appl. Microbiol. Biotechnol*. **93**, 1843-1852.

Lepage, P., Leclerc, M.C., Joossens, M., Mondot, S., Blottière, H.M., Raes, J., Ehrlich, D. and Doré, J. (2013). A metagenomic insight into our gut's microbiome. *Gut*. **62**, 146-158.

Levy, R., and Borenstein, E. (2014). Metagenomics systems biology and metabolic modeling of the human microbiome. *Gut Microbes*. **5**, 1-6.

Ludwig, W. (2007). Nucleic acid techniques in bacterial systematics and identification. *Int.J. Food Microbiol*. **120**, 225-236.

Maiden, M.C.J., Bygraves, J.A., Feil, E., Morelli, G., Russell, J.E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D.A., Feavers, I.M., Achtman, M. and Spratt, B.G. (1998). Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc. Natl. Acad. Sci. USA*. **95**, 3140-3145.

Marean, C.W. (2010). Pinnacle Point Cave 13B (Western Cape Province, South Africa) in context: The Cape Floral kingdom, shellfish, and modern human origins. *J. Hum. Evol*. **59**, 425-443.

Meier-Kolthoff, J. P., Göker, M., Spröer, C. and Klenk, H. P. (2013). When should a DDH experiment be mandatory in microbial taxonomy? *Archiv.Microbiol.* **195**, 413-418.

Park, Y. H., Yoon, J. H., Shin, Y. K., Suzuki, K., Kudo, T., Seino, A., Kim, H., Lee, J. S. & Lee, S. T. (1999). Classification of '*Nocardioides fulvus*' IFO 14399 and *Nocardioides* sp. ATCC 39419 in *Kribbella* gen. nov., as *Kribbella flavida* sp. nov. and *Kribbella sandramycini* sp. nov. *Int J Syst Bacteriol.* **49**, 743-752.

Parte, A. C. (2014). List of prokaryotic names with standing in nomenclature. (http://www.bacterio.net/kribbella.html) Accessed May 2014.

Prescott, L. M., Harley, J. P. & Klein, D. A. (2002). Microbial taxonomy. In: *Microbiology*, 5th Edition, 421-449. USA: The McGraw-Hill Companies, Inc.

Qin, S., Li, J., Chen, H.-H., Zhao, G.-Z., Zhu, W.-Y., Jiang, C.-L., Xu, L.-H. and Li, W.-J. (2009). Isolation, Diversity, and Antimicrobial Activity of Rare Actinobacteria from Medicinal Plants of Tropical Rain Forests in Xishuangbanna, China. *Appl. Environ Microbiol.* **75**, 6176-6186.

Qin, S., Xing, K., Jiang, J.-H., Xu, L.-H.and Li, W.-J. (2011) Biodiversity, bioactive natural products and biotechnological potential of plant-associated endophytic actinobacteria. *Appl. Microbiol. Biot*. **89**, 457-473.

Raes, J. (2014). The gut microbiome- a new target for understanding, diagnosing and treating disease. *Arch. Public Health.* **72,** Suppl 1-K3.

Rajendhran, J. and Gunasekaran, P. (2011). Microbial phylogeny and diversity: Small subunit ribosomal RNA sequence analysis and beyond. *Microbial Research*. **166**, 99-110.

Rebelo, A.G. and Siegfried, W.R. (1990). Protection of Fynbos Vegetation: Ideal and Real-World Options. *Biol.Conserv*. **54,** 15-31.

Richter, M. and Róssello-Móra, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *P. Natl. Acad. Sci .USA.* **106**, 19126-19131.

Riesenfeld, C.S., Schloss, P.D., and Handelsman, J. (2004). Metagenomics: Genomic Analysis of Microbial Communities. *Annu. Rev. Genet*. **38**, 525-552.

Rondon, M.R., August, P.R., Betterman, A.D., Brady, S.F., Grossman, T.H., Liles, M.R., Loiacono, K.A., Lynch, B.A., MacNeil, I.A., Minor, C., Tiong, C.L., Gilman, M., Osburne, M.S., Clardy,J., Handelsman, J. and Goodman, R.M. (2000). Cloning the Soil Metagenome: a Strategy for Accessing the Genetic and Functional Diversity of Uncultured Microorganism. *Appl. Environ. Microbiol*. **66**, 2541-2547.

Rosenblueth, M. and Martínez-Romero, E. (2006). Bacterial Endophytes and Their Interactions with Hosts. *Mol. Plant. Microbe. Interact.* **19**, 827-837.

Santos, S.R. and Ochman, H. (2004). Identification and phylogenetic sorting of bacterial lineages with universally conserved genes and proteins. *Environ. Microbiol*. **6**, 754-759.

Schloss, P.D. and Handelsman, J. (2003). Biotechnological prospects from metagenomics. *Curr. Opin. Biotech*. **14**, 303-310.

Schmeisser, C., Steel, H. andStreit, W.R. (2007). Metagenomics, biotechnology with non-culturable microbes. *Appl. Microbiol. Biot*. **75**, 955-962.

Sohn, K., Hong, S.G., Bae, K.S. and Chun, J. (2003). Transfer of *Hongia koreensis* Lee *et al*. 2000 to the genus *Kribbella* Park *et al.* 1999 as *Kribbella koreensis* comb.nov. *Int. J. Syst. Evol. Microbiol.* **53**, 1005-1007.

Stackebrandt, E. and Ebers, J. (2006). Taxonomic parameters revisited: tarnished gold standards. *Microbiology Today.* November, 152-155.

Stackebrandt, E., Rainey, F. A. and Ward-Rainey, N. L. (1997). Proposal for a new hierarchic classification system, *Actinobacteria* classis nov. *Int. J. Syst. Bacteriol .***47**, 479-491.

Staley, J.T. (2006). The bacterial species dilemma and the genomic-phylogenetic species concept. *Phil. Trans. R. Soc. B*. **361**, 1899-1909.

Staley, J.T. (2009). The phylogenomic species concept. *Microbiol. Today*. **36**, 80-83.

Staley, J.T. (2010). Comprehending Microbial diversity: the fourth goal of microbial taxonomy. *The Bulletin of BISMiS*. **1,** (http://www.bergeys.org/bismisbulletintoc.html).

Strobel, G., Daisy, B., Castillo, U., and Harper, J. (2004). Natural products from endophytic microorganisms. *J. Nat. Prod*. **67**, 257-268.

Sutcliffe, I.C., Trujillo, M.E. and Goodfellow M. (2012). A call to arms for systematists: revitalising the purpose and practises underpinning the description of novel microbial taxa. *Antonie van Leeuwenhoek.* **101**, 13-20.

Taylor, J.E., Lee, S. and Crous, P.W. (2001). Biodiversity in the Cape Floral Kingdom: fungi occurring on *Proteaceae. Mycol. Res*. **105**, 1480-1484.

Tindall, B.J., Rosselló-Móra, R., Busse H.-J., Ludwig, W. and Kämpfer P. (2010). Notes on the characterization of prokaryote strains for taxonomic purposes. *Int. J. Syst. Evol. Microbiol*. **60**, 249-266.

Trujillo, M.E., Kroppenstedt, R.M., Schumann, P., and Martinez-Molina, E. (2006). *Kribbella lupini* sp. nov., isolated from the roots of *Lupinus angustifolius*. *Int. J. Syst. Evol. Microbiol*. **56**, 407-411.

Turnbaugh, P.J., Ridaura, V.K., Faith, J.J., Rey, F.E., Knight, R. and Gordon J.I. (2009). The Effect of Diet on the Human Gut Microbiome: A Metagenomic Analysis in Humanized Gnotobiotic Mice. *Sci. Transl. Med*. **1**, 6ra14.

Urzì, C., De Leo, F. and Schumann, P. (2008). *Kribbella catacumbae* sp. nov. and *Kribbella sancticallisti* sp. nov., isolated from whitish-grey patinas in the catacombs of St. Callistus in Rome, Italy. *Int. J. Syst. Evol. Microbiol.* **58**, 2090-2097.

Ventura, M., Canchaya, C., Tauch, A., Chandra, G., Fitzgerald, G.F., Chater, K.F. and van Sinderen, D. (2007). Genomics of Actinobacteria: Tracing the Evolutionary History of an Ancient Phylum. *Microbiol. Mol. Biol. R*. **71**, 495-548.

Wayne,L.G., Brenner,D.J., Colwell, R.R., Grimont, P.A.D., Kandler, O., Krichevsky, M.I., Moore, L.H., Moore, W.E.C., Murray, R.G.E., Stackebrandt, E., Starr, M.P. and Trüper, H.G. (1987). Report of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics. *Int. J. Syst. Evol. Microbiol.* **37**, 463-364.

Zeigler, D.R. (2003). Gene sequences useful for predicting relatedness of whole genomes in bacteria. *Int. J. Syst. Evol. Microbiol*. **53**, 1893-1900.

Zhi, X.-Y., Li, W.-J. and Stackebrandt, E. (2009). An update of the structure and 16S rRNA gene sequence-based definition of higher ranks of the class *Actinobacteria*, with the proposal of two new suborders and four new families and emended descriptions of the existing higher taxa. *Int. J. Syst. Evol. Microbiol.* **59**, 589–608.

Chapter 2:
A targeted search for *Kribbella* strains from soil and plant samples using culture-dependent methods

Chapter 2: A targeted search for *Kribbella* strains from soil and plant samples using culture dependent methods

Chapter 2: A targeted search for *Kribbella* strains from soil and plant samples using culture-dependent methods

2.1 Summary

Culture-dependent methods were used to isolate *Kribbella* strains from a soil sample, as well as leaf and root samples from 5 indigenous plants. To enhance the chances of isolating this rare actinobacterium, the development and use of a selective *Kribbella* medium was employed. Other media used were International *Streptomyces* Project medium No. 2, Middlebrook 7H9 and a modified SM1, all containing nalidixic acid and cycloheximide. The leaf and root samples were surface sterilized before processing to isolate endophytes. All plates were incubated at 30°C (with a duplicate set at room temperature) for 12 weeks and any colonies which had a *Kribbella*-like morphology were sub-cultured for further study. The soil sample produced 6 strains identified as *Kribbella*. Further testing and sequence analysis using the 16S rRNA gene and *gyrB* gene, indicated that all 6 strains were identical to the type strain of *Kribbella solani*. The plant samples did not produce any colonies identified as *Kribbella*. The *Kribbella* selective medium, although not successful in isolating *Kribbella* strains from the plant samples, was able to enhance the selection of *Kribbella* from a soil sample during an Honours project. Thus, it is suggested that the *Kribbella* selective medium be evaluated further by including it in future *Kribbella* isolation projects.

2.2 Introduction

The search for rare or less common actinomycetes has gained in popularity as the need for new natural products becomes a pressing issue. Filamentous actinobacteria (actinomycetes) are known for their capacity to produce a variety of natural products, ranging from antitumour agents to antimicrobials [Chadhaury *et al.,* 2013; Challis, 2014]. Of all known drugs, 70% have been isolated from actinomycetes and one genus, *Streptomyces*, is responsible for producing 80% of these compounds [Chaudhary *et al.,* 2013; Kurtböke, 2012]. However, with the increasing problem of antibiotic resistance there is a pressing need to look for new antibiotics with new mechanisms of action [Chaudhary *et al.,* 2013]. In an effort to find novel compounds with novel chemical structures, the search has been focussed on overlooked environmental niches and overlooked strains of this beneficial group of bacteria [Qin *et al*., 2009; Tanvir *et al.,* 2014]. As an example, the compound salinisporamide A, a potent cytotoxic proteasome inhibitor, is a promising new antitumor candidate in Phase I clinical trials and was isolated from the rare marine actinomycete *Salinispora tropica*, [Ahmed *et al*., 2013; Fenical *et al*., 2009; Freel *et al*., 2011; Palomo *et al*., 2013]. Rare actinomycetes have been found to produce diverse and complex compounds and in order to tap into this source of natural products a better understanding of their ecology is needed. To do this, isolation methods for these strains need to be improved [Kurtböke, 2012].

Kribbella is one of those genera that are considered rare within the actinomycetes. It currently comprises of 20 species with validly published names [Parte, 2014] and is a relatively young genus, having been established in 1999 by Park *et al.* [1999]. The genus belongs to the family *Nocardioidaceae* (sub-order *Propionibacterineae*) [Park *et al*., 1999; Parte, 2014]. The vegetative mycelium is extensively branched and may fragment into small rod-shaped elements in liquid cultures [Park *et al*., 1999].

Kribbella colonies are pasty with lichenous wrinkled shapes with irregular edges. The predominant menaquinone is MK-9(H_4) and the genomic DNA G+C content ranges from 66-77 mol % [Everest *et al*., 2013; Park *et al*., 1999; Sohn *et al*., 2003]. The diamino

acid in the peptidoglycan is LL-diaminopimelic acid and the major cellular fatty acid is anteiso-$C_{15:0}$ [Park *et al*., 1999]. The major polar lipids are phosphatidylcholine and phosphatidylglycerol. There are no diagnostic sugars present in the whole cell sugar pattern [Everest *et* al., 2013]. Most *Kribbella* species have been isolated from terrestrial soil with *Kribbella hippodromi*, *Kribbella swartbergensis* and *Kribbella karoonensis* isolated from South African soil [Everest and Meyers, 2008; Kirby *et al*., 2006;]. However, some *Kribbella* species have been isolated from various other sources, such as the walls of a catacomb (*Kribbella catacumbae, Kribbella sancticallisti, Kribbella albertanoniae* and *Kribbella italica*) [Urzí *et al*., 2008; Everest *et al*., 2013; Everest *et al*., 2015], and from rock surfaces in an alum slate mine (*Kribbella aluminosa*) [Carlsohn *et al*., 2007]. Interestingly, *Kribbella* species have also been isolated from an overlooked and understudied environmental niche, i.e. living plants. These endophytic species include *Kribbella lupini* isolated from the roots of the plant *Lupinus angustifolius* [Trujillo *et al.,* 2006], *Kribbella endophytica* from the leaves of the tree *Pittosporum angustifolium* [Kaewkla and Franco, 2013] and '*Kribbella podocarpi'*, which was isolated from a South African Yellowwood tree, *Podocarpus latifolius*, by a previous Masters student in the laboratory [Curtis *et al*., manuscript under revision].

Novel natural products associated with plants have been found to be produced by actinomycetes living within the tissues of the plants as endophytes [Castillo *et al*., 2002; El-Shatoury *et al*., 2006; Ezra *et al.,* 2004; Kandpal *et al.,* 2012; Shenpagam *et al.,* 2012]. Endophytes are bacteria cells that live in the inner tissue of living plants and do not visibly harm the plant [Qin *et al.,* 2011]. It has been found that these endophytes are capable of producing an array of secondary metabolites and are an under-explored reservoir of novel actinobacterial species [Qin *et al.,* 2011]. A prime example is the discovery of the peptide antibiotic complex coronamycin. It was found to be produced by an endophytic strain of *Streptomyces* residing in a vine of the genus *Monstera* from the Amazon. Coronomycin has been shown to have activity against fungal pathogens and the malaria parasite *Plasmodium falciparum* [Ezra *et al.,* 2004].

The discovery of coronomycin, from the upper Amazon of Peru, also serves as an example for seeking out areas that have high plant biodiversity [Ezra *et al.,* 2004]. For it is believed that plants growing in areas of great biodiversity have the prospect of also harbouring endophytes with great biodiversity [Strobel, 2003]. The Cape Floral kingdom

(illustrated in Figure 2.1) is, for its size (90,000km^2), the most diverse plant kingdom and has the world's highest representation of endemic plants with an estimated 9030 vascular plants species (68.7% endemic) [Goldblatt and Manning, 2002; Marean, 2010; Taylor *et al.,* 2001]. The Cape Floral kingdom is, therefore, an ideal location to obtain a diverse set of plants to look for novel endophytic strains of actinomycetes.

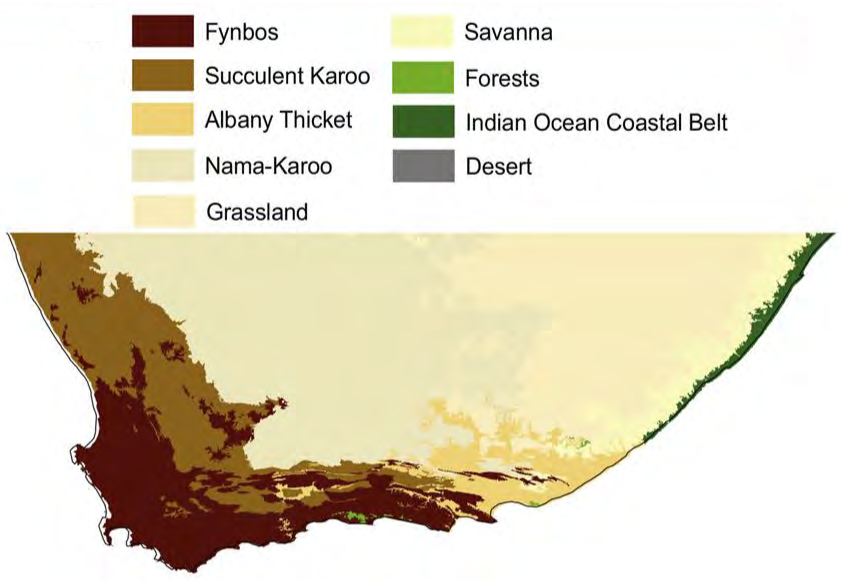


Figure 2.1 Cape Floral Kingdom at the Southern tip of Africa, outlining various vegetative regions [Marean 2010].

For this study, 5 indigenous fynbos plants from the Cape Floral Kingdom were investigated for the presence of *Kribbella* strains living in leaf and root samples along with a soil sample, using plate cultivation techniques. The fynbos vegetation region is represented as a dark brown colour in Figure 2.1. To further enhance the search for *Kribbella*, a selective medium was designed to favour the growth of *Kribbella* strains. *Kribbella* is an ideal candidate to search for since it has been isolated from a wide variety of habitats and, although not a well-known producer of antibiotics, could be a source of novel secondary metabolites.

2.3 Methods and Materials

2.3.1 Sample collection

The soil sample was obtained from the upper Tokai forest (Cape Town, South Africa), approximately 2 meters from the side of the trail. The surface debris and 1cm of top soil were removed before the sample was dug out with a sterile spoon and placed in sterile 50-ml Falcon tube. The soil was stored at ambient room temperature for 2 days before processing.

Five plants were purchased from Kirstenbosch Botanical Gardens nursery as sampling the roots would likely kill any wild plants and disrupt the natural ecology. The plants were, *Agathosma ovata* 'Glentana' (Family *Rutaceae*) pictured in Figure 2.2a, *Orphium frutescens* (Family *Gentianaceae*) Figure 2.2b, *Leucospermum conocarpodendron* (Family *Proteaceae*) Figure 2.3c, *Podylaria sericea* (Family *Fabaceae*) Figure 2.2d and *Protea aurea* subspecies *aurea* (Family *Proteaceae*) Figure 2.2e. The plant samples were kept in the laboratory and watered with sterile de-ionised water until the time of sampling.

Fig.2.2a*) Agathosma ovata* 'Glentana'

Fig.2.2b) *Orphium frutescens*

Fig.2.2c) *Leucospermum conocarpodendron*

Fig.2.2d) *Podylaria sericea*

Fig.2.2e) *Protea aurea* subsp. *aurea*

Figure 2.2a-e. Photographs of the indigenous fynbos plant samples a) *Agathosma ovata* 'Glenatana' [Elands Nursery, 2014], b) *Orphium frutescens* [iSpot, 2014], c) *Leucospermum conocarpodendron* [Notten and January, 2009], d) *Podylaria sericea* [Viljoen, 2007], and e) *Protea aurea* subsp. *aurea* [University of Connecticut, 2014].

A. ovata 'Glentana' grows in the Western Cape Province as an upright evergreen shrub to a height of 1.5 m (Figure 2.3a) [Jodamus, 2001]. *O. frutescens* is found in the south-western Western Cape Province growing as a bushy perennial to a height of 60 cm (Figure 2.3b). The leaves and stems have small white hairs and the plant is able to grow close to the sea. *O. frutescens* is the only species in the genus *Orphium* [Van der Walt, 2000]. *L. conocarpodendron* is a tree-like shrub reaching 3-5 meters tall. The leaves are grey-green and covered in a dense layer of fine hairs. *L. conocarpodendron* is endemic to the Cape Peninsula, Cape Town, from Devil's Peak and Table Mountain to the Twelve Apostles and Llandudno (Figure 2.3c) [Notten and January, 2009]. *P. sericea* is a rounded evergreen shrub reaching 1 meter in height. It is endemic to southern Africa, mainly occurring in the southern Western Cape Province (Figure 2.3d). The leaves are long and straight with glossy hairs above and beneath which give the plant a silver-grey lustre [Viljoen, 2007]. *P. aurea* subspecies *aurea* is a small tree reaching up to 4 meters tall. The leaves are grey-green and are soft and velvety when young. It is endemic to the Cape Floristic region (Figure 2.3e) [Goldblatt and Manning, 2000; Nurrish, 2009].

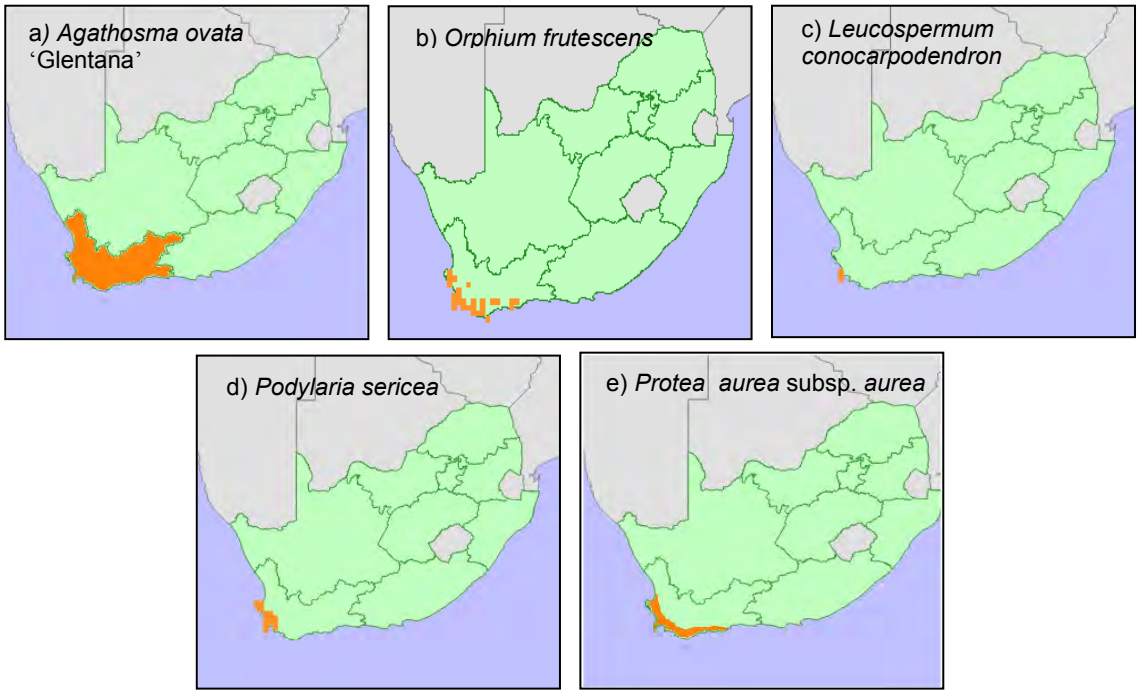


Figure 2.3 a-e Natural ranges of the indigenous plants indicated in orange within the country of South Africa which is highlighted in green. a) *Agathosma ovata* 'Glenatana' adapted from [Schutte-Vlok, and Raimondo, 2012], b) *Orphium frutescens* [Foden and Potter, 2005], c) *Leucospermum conocarpodendron* [Rebelo *et al*., 2005], d) *Podylaria sericea* [Schutte-Vlok, and Raimondo, 2012], and e) *Protea aurea* subsp. *aurea* adapted from [Schutte-Vlok, and Raimondo, 2012].

2.3.2 Kribbella selective medium (KSM)

Kribbella strains are less frequently isolated in culture-dependant laboratory studies compared to other genera, such as *Streptomyces*. In order to increase the chances of isolating *Kribbella* strains, a selective minimal medium was developed. The minimal medium recipe was based on the basal mineral salts recipe of the medium used to test the utilization of sole nitrogen sources [Locci, 1989; Shirling and Gottlieb, 1966]. To this, specific carbon and nitrogen sources were added. A survey of all the papers describing *Kribbella* species was undertaken to identify carbon- and nitrogen-source utilization patterns. Particular attention was paid to identifying carbon and nitrogen sources that are not commonly utilized by actinobacteria. This analysis revealed that all *Kribbella* species can utilize adonitol as a carbon source and 4-L-hydroxyproline as a nitrogen source. The resulting medium, named *Kribbella* selective medium (KSM), is formulated below. The basal-medium agar and adonitol solutions are prepared and sterilized

separately. Once cooled to 60°C the two components are combined before pouring plates.

Recipe for 1 Litre KSM media

KH_2PO_4	2.38g
K_2HPO_4 $3H_2O$	5.65g
$MgSO_4$ $7H_2O$	1.0g
Trace salts solution*	1.0 ml
4-L-Hydroxyproline	2.0g

Water to 900ml
pH – 7.0
Agar – 15g

Adonitol – 10g in 100ml water autoclaved separately

*Trace salts solution consists of

$FeSO_4$ $7H_2O$	0.1g
$MnCl_2$ $4H_2O$	0.1g
$ZnSO_4$ $7H_2O$	0.1g

Water to 100ml
Sterilize by autoclaving

2.3.3 Leaf and root surface sterilization procedure and sample preparation

The following surface sterilization procedure used for the leaf and root samples was based on the procedures of Coombs & Franco [2003] and Kirby & Meyers [2010] with the additional steps of ultraviolet radiation exposure and multiple rinses with sterile de-ionised water.

Sampled roots were rinsed under running water to dislodge soil.
Air dried the leaf/root samples overnight at room temperature
Sonicated for 2 minutes in 70% ethanol (EtOH)
1 minute exposure to 254 nm ultraviolet radiation on one side, then flipped over and exposed for another minute
Washed in 100% EtOH for 5 minutes, then rinsed with sterile de-ionised water for 1 minute
Washed in 3.5% (w/v) NaOCl for 5 minutes, then rinsed with sterile de-ionised water for 1 minute
Washed in 70% EtOH for 5 minutes, rinsed with sterile de-ionised water for 1 minute
Followed by 2 more 1 minute rinses in sterile de-ionised water.

After the procedure, the samples were dabbed onto fresh International *Streptomyces* Project medium No. 2 (ISP2) plates before further processing, to confirm that the sterilization procedure was successful. The sterilization-test plates were incubated at 30°C for 7 days.

2.3.4 Plate isolations

Soil sample

The soil was first pretreated at 60°C for 1 hour to decrease the amount of vegetative bacteria. 0.1g of the pretreated soil sample was added to 1ml of sterile de-ionised water and vortexed for 1 minute. The sample was then serially diluted with sterile water to 10^{-4}. All the dilutions were plated on ISP2, Difco Middlebrook 7H9 agar containing 10mM of glucose (albumin-catalase supplement omitted; Becton Dickinson), Czapek solution agar (CZ) [Atlas, 2004]; and a modified version of Selective Medium 1 (SM1) [Tan *et al.*, 2006]. The modified SM1 differed by the use of 5.58g of yeast extract and 0.416g of ammonium sulphate instead of 6.7g of yeast nitrogen base and by omitting the nystatin [Tan *et al*., 2006]. The plates were then incubated in sealed bags with moist paper towels at 30°C for 12 weeks.

Plant samples

Approximately 0.5g of root material was sampled from each plant along with approximately 2.0g of leaf material. After surface sterilization, the leaf and root sections were sterilely cut into < 5mm slices and 4-5 slices of leaf or root were placed directly on to sterile plates of ISP2, 7H9, modified SM1 and KSM, in duplicate. The remainder of the sliced material was incubated at 30°C with shaking for 1 hour in a 10ml quarter strength phosphate buffer solution [Kirby and Meyers, 2010]. This suspension was spread plated on the previously mentioned media (ISP2, KSM, 7H9 and modified SM1) at undiluted, 1/2 and 1/5-diluted with quarter strength phosphate buffer. The plates were then incubated at 30°C for 12 weeks for the *Agathosma* plant samples and at 30°C and room temperature for 12 weeks for the remaining plant samples.

All media used in the isolation procedures contained 10µg/ml nalidixic acid and 50µg/ml cycloheximide to deter the growth of Gram negative bacteria and fungi, respectively. Plates were periodically inspected for beige/cream wrinkled and convoluted colonies resembling the morphology of *Kribbella.* Any colonies resembling *Kribbella* were sub-cultured onto fresh plates of the same medium from which they were isolated (without antibiotics), for further study.

2.3.5 Genomic DNA extraction

Suspected *Kribbella* isolates were grown in 20ml of liquid ISP2 broth in 250ml Erlenmeyer flasks. Approximately 200µl of cell mass was obtained by centrifuging the liquid culture in 1.5ml benchtop centrifuge tubes. The genomic DNA was extracted using a modified version of the phenol extraction method of Wang *et al.*, [1996], increasing the lysozyme concentration to 20mg/ml and the digestion period to overnight [Everest *et al.,* 2011].

2.3.6 16S rRNA gene and gyrB gene amplification

The extracted DNA was then amplified by polymerase chain reaction (PCR). The universal 16S rRNA gene primers F1 and R5 were used in a standard PCR program to amplify 16S rRNA gene sequences [Cook and Meyers, 2003]. The *gyrB* gene primers KgyrB953F and KgyrB-R (KR) were used to amplify a short variable region of the *gyrB* gene [Kirby *et al*., 2010]. All PCR reactions were carried out in 50µl volumes and contained 2 mM $MgCl_2$, 0.5U SuperTherm *Taq* polymerase (JMR Holdings, USA), 150 µM of each dNTP, 0.5µM of each primer and 500ng template DNA. PCR was performed using a Techne TC-512 thermal cycler. The standard PCR program followed an initial denaturation at 96°C for 2 min, 30 cycles of denaturation at 96°C for 45s, annealing at 56°C for 30s and extension at 72°C for 1 minute 30s, with a final extension at 72°C for 5 minutes. The resulting PCR products were purified using an MSB® Spin PCRapace kit (Invitek, Berlin, Germany). Sequencing was performed as a service by Macrogen Inc., Seoul, South Korea. The sequences were edited manually with Chromas version 2.01 (Technelysium Pty. Ltd., Australia) and assembled with DNAMAN version 5.2.9 (Lynnon BioSoft). The sequences were analyzed by BLAST against the GenBank database and, in the case of the 16S rRNA gene sequences, additionally with the EzTaxon-e database [Kim *et al*., 2012], selecting the search option to include both cultured and uncultured strains.

2.3.7 Preliminary identification of soil isolates by the Cook and Meyers method

The PCR-amplified 16S rRNA gene sequences were used to quickly identify the isolates to the genus level. The method developed by Cook & Meyers [2003], utilizes single

restriction endonuclease digestion banding patterns to identity the strain to the genus level. Each restriction endonuclease reaction contained 5-10µl of 16S rRNA gene PCR product, 2U of the restriction endonuclease, 2 µl of the appropriate buffer and sterile water to give a final volume of 20 µl. The reaction was incubated overnight in a 37°C waterbath, then mixed with 2 µl of 6x tracking dye and electrophoresed on a 0.8% agarose gel along with an undigested 16S rRNA gene PCR product control and λ-*Pst*I molecular weight marker [Cook and Meyers, 2003; Sambrook *et al*., 1989]. The resulting banding pattern and size of the band(s) dictates the next restriction enzyme to use in the method until a genus level identification is obtained. For the colonies tested here, the following restriction endonucleases were used, *Mbo*I (isoschizomer of *Sau3*AI), *Vsp*I (isoschizomer of *Asn*I), *Sph*I, *SnaB*I, *Sal*I, *Age*I and *Sts*I.

2.3.8 Antibiotic screening

To test the isolates for antibiotic activity an overlay plate technique was utilized. Isolated strains were tested against *Escherichia coli* ATCC 25922, *Staphylococcus aureus* ATCC 25923 and *Mycobacterium aurum* A+. The isolates were stab inoculated onto ISP2 plates and tested after 7 days incubation at 30°C. The *E. coli, S. aureus* and *M. aurum* strains were each inoculated into 10 ml of liquid Luria-Bertani [Sambrook *et al*., 1989] medium and incubated at 37°C overnight with shaking. To standardize the amount of cells used in the test, the grown cultures of *E. coli, S. aureus* and *M. aurum* had their optical density at 600nm (OD_{600}) measured by spectrophotometer. The optical density reading was then used to determine the volume of culture, in microlitres, to add to each plate using the following equation OD_{600} x µl = 160 for *M. aurum* and *S. aureus* and OD_{600} x µl = 4 for *E. coli*. The determined amount of culture was inoculated into 6ml volumes of sterile Luria-Bertani sloppy agar (0.7% agar) in glass test tubes [Sambrook *et al*., 1989]. The sloppy agar containing the test organisms was then poured onto the stab inoculated plates. The overlay plates were incubated at 37°C overnight for *E. coli* and *S. aureus* and 48 hours for *M. aurum*. Zones of growth inhibition were then measured and the area of inhibition calculated.

2.4 Results

2.4.1 Kribbella Selective Medium (KSM)

All *Kribbella* type strains were inoculated onto KSM and incubated at 30°C for 7 days to determine whether they could grow on this medium. The KSM medium was also tested using 9 strains from 4 other genera (*Gordonia*, *Amycolatopsis, Actinomadura* and *Streptomyces*) to see if they would grow. It was found that all of the *Kribbella* type strains grew on KSM, but that the majority of the non-*Kribbella* strains could not grow on KSM. However, one *Streptomyces* species (strain M61), two *Amycolatopsis* strains (*Amycolatopsis endophytica* KCTC 19776^T, and M29) and one *Actinomadura* strain M27 were able to grow very weakly on the *Kribbella* selective medium over 7 days at 30°C. Of the various renditions of a medium to select for the growth of *Kribbella* strains, the adonitol and 4-L-hydoxyproline combination (KSM) allowed for growth of all *Kribbella* type strains with the least amount of growth of the other genera tested. The use of KSM in an Honours student's project was very successful and produced 5 of the 10 strains of *Kribbella* isolated from a soil sample using four different isolation media [Mbobo, 2013]. As KSM was developed relatively late in this Ph.D. study, it was not used in the isolation of *Kribbella* strains from soil.

2.4.2 Kribbella isolations from soil

From the soil isolation plates, five colonies matching the colony morphology of *Kribbella* were chosen from ISP2 plates and one colony from a 7H9 plate.

From preliminary identification using the Cook and Meyers method [2003], the six chosen isolates (strains PY4, SY3, SY8, SY10, SH5 and SY9) proved to have restriction endonuclease banding patterns indicative of the genus *Kribbella* and were thus presumptively identified as *Kribbella* strains.

2.4.3 Sequencing results

For all six isolates, the 16S rRNA gene sequence top hit was identical in the EzTaxon-e (*K. hippodromi* strain S1.4^T) and GenBank (*Kribbella solani* strain YB2) databases. The top hit for each database may be seen in Table 2.1. When all 16S rRNA sequences were aligned in MEGA 5.0 all six sequences were identical in the overlapping region.

Table 2.1. Sequencing top hits for the EzTaxon-e and GenBank databases based on 698 bp of 16S rRNA gene sequence and 245 bp of *gyrB* sequence

Soil isolate	16S EzTaxon-e database	% Similarity	16S GenBank top hit	% Similarity	*gyrB* GenBank top hit	% Similarity
SY3	*Kribbella hippodromi* S1.4(T)	100	*Kribbella solani* YB2	100		
PY4	*Kribbella hippodromi* S1.4(T)	100	*Kribbella solani* YB2	100	*Kribbella solani* CIP 108508(T)	100
SY8	*Kribbella hippodromi* S1.4(T)	100	*Kribbella solani* YB2	100		
SY10	*Kribbella hippodromi* S1.4(T)	100	*Kribbella solani* YB2	100		
SH5	*Kribbella hippodromi* S1.4(T)	100	*Kribbella solani* YB2	99	*Kribbella solani* CIP 108508(T)	100
SY9	*Kribbella hippodromi* S1.4(T)	100	*Kribbella solani* YB2	99	*Kribbella solani* CIP 108508(T)	100

The type strains of *K. hippodromi* and *K. solani,* like many *Kribbella* type strains, share a high 16S rRNA gene sequence similarity. *K. hippodromi* and *K. solani* share 99.57% 16S rRNA gene similarity, but have a DNA-DNA relatedness value of 40.4 ±3.8%, indicating that they are distinct species [Everest and Meyers, 2008]. The differing 16S rRNA gene top hits obtained from the EzTaxon-e and GenBank databases are probably due to the use of different sequence alignment algorithms. However, *K. hippodromi* and *K. solani* are very closely related species. Due to the discrepancy in the results from the two databases, a 245 bp portion of the *gyrB* gene was amplified and sequenced to further identify the isolates PY4, SH5 and SY9. Strain PY4 was chosen to represent the group made up of PY4, SY3, SY8 and SY10, since the 16S rRNA gene sequencing results were identical to each other. The region of *gyrB* amplified was within the variable region noted by Kirby *et al*. [2010]. The *gyrB* gene sequences for all three strains also returned the same top hit (*K. solani*) when analyzed using the GenBank database. The top hits can be seen in Table 2.1. Consistent with the 16S rRNA gene sequence results, when aligned the *gyrB* sequences were identical. The genetic distance values between the six soil isolates and all *Kribbella* type strains were calculated [Kirby *et al*., 2010] for the *gyrB* gene sequences (Table 2.2). The soil isolates were found to have a value of 0.000 between each other and against both *K. solani* CIP108508^T and *K. solani* strain YB2. These values are highlighted in blue in Table 2.2. A *gyrB* genetic distance of 0.000 indicates that the strains belong to the same species (in all cases examined, the type strain and all non-type strains of a *Kribbella* species have identical *gyrB* sequences; Kirby *et al*., 2010; Curtis and Meyers, 2012). From Table 2.2, the *gyrB* genetic distance between strains PY4, SH5 and SY9 and *K. hippodromi* DSM 19227^T was 0.025 (highlighted in pink). This suggests that these three soil isolates do not belong to the

species *K. hippodromi*, despite the high similarity of the 16S rRNA gene sequencing results from EzTaxon-e. Thus, using the *gyrB* gene sequence-similarity and genetic distance data, it was concluded that all six isolated strains from the soil sample are identical and belong to the species *K. solani*.

Table 2.2. *gyrB* genetic distance values for soil isolates against all type strains in the genus *Kribbella* based on the 245-nt section within the 315 nt *gyrB* variable region [Kirby *et al*., 2010].

	***Kribbella* strains**	**PY4**	**SH5**	**SY9**
1	**PY4**			
2	**SH5**	**0.000**		
3	**SY9**	**0.000**	**0.000**	
4	*Kribbella alba* DSM 15500^T	0.115	0.115	0.115
5	*Kribbella aluminosa* DSM 18824^T	0.042	0.042	0.042
6	*Kribbella aluminosa* HKI 0480	0.042	0.042	0.042
7	*Kribbella aluminosa* HKI 0479	0.042	0.042	0.042
8	*Kribbella amoyensis* XMU198^T	0.073	0.073	0.073
9	*Kribbella antibiotica* DSM 15501^T	0.008	0.008	0.008
10	*Kribbella albertanoniae* BC640^T	0.096	0.096	0.096
11	*Kribbella catacumbae* DSM 19601^T	0.105	0.105	0.105
12	*Kribbella catacumbae* BC628	0.105	0.105	0.105
13	*Kribbella catacumbae* BC630	0.105	0.105	0.105
14	*Kribbella catacumbae* BC632	0.105	0.105	0.105
15	*Kribbella endophytica* DSM 23718^T	0.105	0.105	0.105
16	*Kribbella italica* BC637^T	0.082	0.082	0.082
17	*Kribbella flavida* DSM 17836^T	0.073	0.073	0.073
18	*Kribbella ginsengisoli* DSM 17941^T	0.110	0.110	0.110
19	***Kribbella hippodromi* DSM 19227^T**	**0.025**	**0.025**	**0.025**
20	*Kribbella jejuensis* CIP 108509^T	0.034	0.034	0.034
21	*Kribbella karoonensis* DSM 17344^T	0.042	0.042	0.042
22	*Kribbella koreensis* CIP 108301^T	0.110	0.110	0.110
23	*Kribbella lupini* DSM 16683^T	0.100	0.100	0.100
24	*Kribbella sancticallisti* DSM 19602^T	0.087	0.087	0.087
25	*Kribbella sancticallisti* BC634	0.087	0.087	0.087
26	*Kribbella sandramycini* DSM 15626^T	0.086	0.086	0.086
27	*Kribbella shirazensis* DSM 45490^T	0.033	0.033	0.033
28	***Kribbella solani* CIP 108508^T**	**0.000**	**0.000**	**0.000**
29	***Kribbella solani* YB2**	**0.000**	**0.000**	**0.000**
30	*Kribbella swartbergensis* DSM 17345^T	0.055	0.055	0.055
31	*Kribbella yunnanensis* DSM 15499^T	0.096	0.096	0.096
32	*Microlunatus phosphovorus* NM-1^T	0.227	0.227	0.227

2.4.4 Antibiotic screening

The three isolates tested were PY4, SH5 and SY9. As strain PY4 had identical 16S rRNA gene sequences to strains SY3, SY8 and SY10, it was chosen to represent the group. All three strains showed no activity against *E. coli* and *S. aureus*. However, all three isolates did show weak activity against *M. aurum*. The areas of the zones of inhibition were PY4 – 176 mm^2; SH5 - 149 mm^2; and SY9 – 163 mm^2.

2.4.5 Plant isolations

Although up to 25 bacterial colonies per plate were obtained, no colonies exhibiting *Kribbella*-like colony morphology were observed on any of the isolation plates (four different media) for the root and leaf samples during the incubation period of 12 weeks.

2.5 Discussion

The aim of this part of the study was to isolate *Kribbella* strains from soil and the leaves and roots of five indigenous plants. The soil isolations were successful in isolating six *Kribbella* strains, however, all six strains proved to be identical to each other by 16S rRNA gene and *gyrB* gene sequencing. Kirby *et al*. [2010] noted that a *gyrB* genetic distance value of 0.04 can be used as a threshold to distinguish between *Kribbella* species and that any genetic distance value under the 0.04 threshold would require DNA-DNA hybridization in order to determine whether two strains belong to the different species. However, there are exceptions. The type strains of *K. hippodromi* and *K. solani* have a *gyrB* genetic distance value of 0.025, but the DNA-DNA hybridization value (40.4%) is well below the threshold of 70% for distinguishing species [Wayne *et al*., 1987]. This pattern can also be seen in the comparison of the type strains of *Kribbella ginsengisoli* and *Kribbella koreensis* (*gyrB* genetic distance for this region: 0.004. DNA relatedness: 45%) [Cui *et al*., 2010]. The *gyrB* genetic distance values for the 245 bp region of the entire genus are shown in Appendix 2A. All values under the 0.04 threshold are highlighted in pink and those highlighted in blue are values of 0.000. It can be seen from the blue-highlighted values that only multiple strains of a species share a *gyrB* genetic distance of 0.000. Using this data, all the isolated strains are identical to each other and, even though the 16S rRNA gene sequencing results from the EzTaxon-e database suggested that *K. hippodromi* was the closest match, the *gyrB* data indicated that the six soil isolates have identical sequences to *K. solani* and are thus strains of this species. Although the isolated strains do not belong to a new species, they did exhibit weak antibacterial activity against *M. aurum*.

The endophyte isolations were unfortunately not successful, which may be due to the smaller sample size of material and the lower number of bacteria present in plant samples than soil samples. Isolations might have improved if larger amounts of leaf and root material were sampled and larger volumes of the extract were screened and/or the sample was concentrated through membrane filters followed by placing the filters onto plates of various media [Slanetz *et al*., 1955]. Soil contains the greatest density and diversity of actinomycetes and may contain up to 10 billion microbes per gram with an estimated species richness of 4000 species per gram of soil [Daniel, 2004; Janso and Carter, 2010; Riesenfeld *et al*., 2004]. Endophytes, however, have a highly variable population density, which is dependant on numerous factors including the type of host plant, growth stage of the host and environmental factors [Senthikumar *et al*., 2011]. One study on potato endophytes found an average endophyte population of 2.2×10^6 CFU/g fresh weight of roots, not including the tubers, and an overall average endophyte population density of 3.9×10^4 CFU/g fresh weight of leaf and root material combined [Berg *et al*., 2005]. Densities of culturable endophytic bacteria are found to be between 10^2 and 10^6 CFU/g of plant material, whereas a single gram of soil can harbour up to 10 billion bacterial cells [Berg *et al*., 2005; Raynaud and Nunan, 2014]. Since there is an estimated factor of $10^4 – 10^8$ cells per gram less in plant material as compared to soil, it may prove beneficial to screen larger amounts of plant sample, in order to increase the effectiveness of isolating endophytic actinobacteria.

The unsuccessful endophyte isolations may also in part be due to various nutrient requirements. Lower nutrient medium (i.e. lower concentrations of nutrients) may prove more effective due to the fact that media with high nutrient concentrations allow fast growing bacteria to out-compete and overgrow the slower growing actinobacteria [Qin *et al*., 2011]. Qin *et al*. [2011] also found that using media with amino acids as nitrogen sources and cellulose, xylan and sodium succinate as carbon sources proved effective, as plant amino acids account for the majority of an endophyte's nitrogen source and cellulose and xylan are components in the plant cell wall. Further attempts to isolate actinobacterial endophytes could include lowering the nutrient content of the media or adding differing amounts of plant extracts into the media in order to increase the effectiveness of the isolations [Qin *et al*., 2011]. Although the KSM medium developed here did not lead to the isolation of *Kribbella* strains in this study, the use of KSM in an Honours student's project was successful and produced half of all the *Kribbella* strains

isolated from a soil sample [Mbobo, 2013]. Therefore, including KSM as an isolation medium in further *Kribbella* isolation projects should prove beneficial.

Given that approximately only 1% of the bacterial diversity in most environmental samples has been cultured using standard culturing techniques [Riesenfeld *et al*., 2004], there is a strong need to improve upon actinobacterial endophyte isolation. The currently uncultured microbes may represent a vast resource for the development of enzymes, antibiotics and other chemical compounds for use in biotechnology [Schmeisser *et al*., 2007]. One important tool to access and investigate this potential reservoir of actinobacteria is metagenomics [Schmeisser *et al*., 2007]. Metagenomics is a culture independent tool that can assess the entire population of microbes in a sample [Culligan *et al*., 2014; Schmeisser *et al*., 2007] and was used to further investigate the indigenous plant samples in this study. This work is presented in Chapter 3.

2.6 References

Ahmed, L., Jensen, P.R., Freel, K.C., Brown, R., Jones, A.L., Kim, B.-Y. and Goodfellow, M. (2013). *Salinispora pacifica* sp. nov., an actinomycete from marine sediments. *Antonie van Leeuwenhoek*. **103**, 1069–1078.

Atlas, R. M. (2004). *Handbook of Microbiological Media (3rd edition)*. Boca Raton, Fl: CRC Press.

Berg, G., Krechel, A., Ditz, M., Sikora, R.A., Ulrich, A. and Hallman, J. (2005). Endophytic and ectophytic potato-associated bacterial communities differ in structure and antagonistic function against plant pathogenic fungi. *FEMS Microbiol Ecol*. **51**, 215-229.

Carlsohn, M.R., Groth, I., Spröer, C., Schütze, B., Saluz, H.-P., Munder, T., and Stackebrandt, E. (2007). *Kribbella aluminosa* sp. nov., isolated from a medieval alum slate mine. *Int. J. Syst. Evol. Microbiol.* **57**, 1943-1947.

Castillo, U. F., Strobel, G. A., Ford, E. J., Hess, W.M., Porter, H., Jensen, J.B., Albert, H., Robison, R., Condron, M.A.M., Teplow, D.B., Stevens, D. and Yaver, D. (2002). Munumbicins, wide spectrum antibiotics produced by *Streptomyces* NRRL 30562, endophytic on *Kennedia nigriscans*. *Microbiology*. **148**, 2675–2685.

Challis, G.L., (2014). Exploitation of the *Streptomyces coelicolor* A3(2) genome sequence for discovery of new natural products and biosynthetic pathways. *J. Int. Microbiol. Biotechnol.* **41**, 219–232.

Chaudhary H.S., Soni, B., Shrivastava, A.R., Shrivastava, S. (2013). Diversity and Versatility of Actinomycetes and its Role in Antibiotic Production. *J. App. Pharm Sci.* **3**, S83-S94.

Cook, A.E. and Meyers, P.R. (2003). Rapid identification of filamentous actinomycetes to the genus level using genus-specific 16S-rRNA gene restriction fragment patterns. *Int. J. Syst. Evol. Microbiol*. **53**, 1907-1915.

Coombs, J.T. and Franco, C.M.M. (2003). Isolation and Identification of Actinobacteria from Surface-Sterilized Wheat Roots. *Appl. Environ. Microbiol*. **69**, 5603-5608.

Culligan, E.P., Sleator, R. D., Marchesi, J. R., and Hill, C. (2014). Metagenomics and novel gene discovery: Promise and potential for novel therapeutics. *Virulence*. **5**, 399-412.

Curtis, S.M., Norton, I., Everest, G.J. and Meyers, P.R. (2015). *Kribbella podocarpi* sp. nov., isolated from the leaves of a yellowwood tree (*Podocarpus latifolius*) and *Kribbella capetownensis* sp. nov. and *Kribbella speibonae* sp. nov., isolated from soil. *Antonie van Leeuwenhoek.* Manuscript under revision.

Daniel, R. (2004). The soil metagenome- a rich resource for the discovery of novel natural products. *Curr. Opin. Biotech*. **15**, 199-204.

Elands Nursery (2014). Products: *Agathosma ovata* 'Glentana'. Accessed May 2014. http://www.elands.co.za/ProductView.aspx?product=1745

El-Shatoury, S., Abdulla, H., El-Karaaly, O., El-Kazzaz, W. and Dewedar, A. (2006). Bioactivies of Endophytic Actinomycetes from Selected Medicinal Plants in the World Heritage Site of Sain Katherine, Egypt. *Int. J. Bot*. **2**, 307-312.

Everest, G.J., Cook, A.E., Kirby, B.M. and Meyers, P.R. (2011). Evaluation of the use of recN sequence analysis in the phylogeny of the genus *Amycolatopsis*. *Antonie van Leeuwenhoek.* **100**, 483-496.

Everest, G.J., Curtis, S.M., De Leo, F., Urzí, C. and Meyers, P.R. (2013). *Kribbella albertanoniae* sp. nov., isolatated from a Roman catacomb, and emended description of the genus *Kribbella*. *Int. J. Syst. Evol. Microbiol*. **63**, 3591-3596.

Everest, G.J., Curtis, S.M., De Leo, F., Urzí, C. and Meyers, P.R. (2015). Description of *Kribbella italica* sp. nov., isolated from a Roman catacomb. *Int. J. Syst. Evol. Microbiol*. **64**, 000-000. *In press*.

Everest, G.J., and Meyers, P.R. (2008). *Kribbella hippodromi* sp. nov., isolated from soil from a racecourse in South Africa. *Int. J. Syst. Evol. Microbiol*. **58**, 443-446.

Ezra, D., Castillo, U.F., Strobel, G.A., Hess, W.M., Porter, H., Jensen, J.B., Condron, M.A.M., Teplow, D.B., Sears, J., Maranta, M., Hunter, M., Weber, B. and Yaver, D. (2004) Coronamycins, peptide antibiotics produced by a verticillate *Streptomyces* sp. (MSU-2110) endophytic on *Monstera* sp. *Microbiology*. **150**, 785-793.

Fenical, W., Jensen, P.R., Palladino, M.A., Lam, K.S., Lloyd, G.K., Potts, B.C. (2009). Discovery and development of the anticancer agent salinosporamide A (NPI-0052). *Bioorg. Med. Chem*. **17**, 2175–2180.

Foden, W. and Potter, L. (2005). *Orphium frutescens (L.)* E.Mey. National Assessment: Red List of South African Plants version 2014.1. Accessed November 2014. http://redlist.sanbi.org/species.php?species=1894-1.

Freel, K.C., Nam, S.-J., Fenical, W., Jensen, P.R. (2011). Evolution of secondary metabolite genes in three closely related marine actinomycete species. *Appl. Environ. Microbiol.* **77**, 7261–7270.

Goldblatt, P. and Manning, J.C. (2000). *Cape Plants: A Conspectus of the Cape Flora of South Africa*. Pretoria, South Africa. National Botanical Institute of South Africa.

Goldblatt, P. and Manning, J.C. (2002). Plant Diversity of the Cape Region of Southern Africa. *Annals of the Missouri Botanical Garden*. **89**, 281-302.

iSpot (2014). Observations in the species: *Orphium frutescens* Accessed May 2014. http://www.ispot.org.za/species_dictionary/Orphium%20frutescens

Janso, J.E. and Carter G.T. (2010). Biosynthetic Potential of Phylogenetically Unique Endophytic Actinomycetes from Tropical Plants. *Appl. Environ. Microbiol*. **76**, 4377-4386.

Jodamus, N. (2001). *Agathosma ovata 'Kluitjieskraal'*. PlantZAfrica.com. South African National Biodiversity Institute. Accessed May 2014. http://www.plantzafrica.com/frames/plantsfram.htm.

Kaewkla, O. and Franco, C.M.M. (2013). *Kribbella endophytica* sp. nov., an endophytic actinobacterium isolated from the surface-sterilized leaf of a native apricot tree. *Int. J. Syst. Evol. Microbiol.* **63**, 1249-1253.

Kandpal, K.C., Jain, D.A., Kumar, U., Tripathi, R. and Kumar, T.S. (2012). Isolation and screening of endophytic actinomycetes producing antibacterial compound from *Citrus aurantifolia* Fruit. *Eur. J. Exp. Biol.* **2**, 1733-1737.

Kim, O.S., Cho, Y.J., Lee, K., Yoon, S.H., Kim, M., Na, H., Park, S.C., Jeon, Y.S., Lee, J.H., Yi, H., Won, S., Chun, J. (2012). Introducing EZ-Taxon: a prokaryotic 16S rRNA Gene sequence database with phylotypes that represent uncultured species. *Int J Syst Evol Microbiol* **62**, 716–721.

Kirby, B.M., Everest, G.J., and Meyers, P.R. (2010). Phylogenetic analysis of the genus *Kribbella* based on the *gyrB* gene: proposal ofa *gyrB*-sequence threshold for species delineation in the genus *Kribbella*. *Antonie van Leeuwenhoek*. **97**, 131-142.

Kirby, B.M., Le Roes, M., Meyers, P.R. (2006). *Kribbella karoonensis* sp. nov. and *Kribbella swartbergensis* sp. nov., isolated from soil from the Western Cape, South Africa. *Int. J. Syst. Evol. Microbiol.* **56**, 1097-1101.

Kirby, B.M. and Meyers, P.R. (2010). *Micromonospora tulbaghiae* sp. nov., isolated from the leaves of wild garlic, *Tulbaghia violacea*. *Int. J. Syst. Evol. Microbiol.* **60**, 1328-1333.

Kurtböke, D. Í. (2012). Biodiscovery from rare actinomycetes: an eco-taxonomical perspective. *Appl. Microbiol. Biotechnol*. **93**, 1843-1852.

Locci, R. (1989). *Streptomyces* and related genera Bergey's Manual of Systematic Bacteriology. Williams and Wilkins Company, Baltimore. **4**, 2451-2508.

Marean, C.W. (2010). Pinnacle Point Cave 13B (Western Cape Province, South Africa) in context: The Cape Floral kingdom, shellfish, and modern human origins. *J Hum Evol*. **59**, 425-443.

Mbobo B. (2013). Selective isolation of *Kribbella* strains from the environment. B.Sc. (Honours) thesis. Department of Molecular and Cell Biology, University of Cape Town. Supervisor Paul Meyers

Notten, A. and January, E. (2009). *Leucospermum conocarpodendron*. PlantZAfrica.com. South African National Biodiversity Institute. Accessed May 2014. http://www.plantzafrica.com/frames/plantsfram.htm

Nurrish, L. (2009). *Protea aurea*. PlantZAfrica.com. South African National Biodiversity Institute. Accessed May 2014. http://www.plantzafrica.com/frames/plantsfram.htm

Palomo, S., González, I., de la Cruz, M., Martín, J., Tormo, J.R., Anderson , M., Hill, R.T., Vicente, F., Reyes, F. and Genilloud, O. (2013). Sponge-Derived *Kocuria* and *Micrococcus* spp. as Sources of the New Thiazolyl Peptide Antibiotic Kocurin. *Mar. Drugs*. **11**, 1071-1086.

Park, Y. H., Yoon, J. H., Shin, Y. K., Suzuki, K., Kudo, T., Seino, A., Kim, H., Lee, J. S. & Lee, S. T. (1999). Classification of '*Nocardioides fulvus*' IFO 14399 and *Nocardioides* sp. ATCC 39419 in *Kribbella* gen. nov., as *Kribbella flavida* sp. nov. and *Kribbella sandramycini* sp. nov. *Int J Syst Bacteriol.* **49**, 743-752.

Parte, A. C. (2014). List of prokaryotic names with standing in nomenclature. Accessed May 2014. http://www.bacterio.net/kribbella.html

Qin, S., Li, J., Chen, H-H., Zhao, G-Z., Zhu, W-Y., Jiang, C-L., Xu, L-H. and Li, W-J. (2009). Isolation, Diveristy, and Antimicrobial Activity of Rare Actinobacteria from Medicinal Plants of Tropical Rain Forests in Xishuangbanna, China. *Appl. Environ Microbiol.* **75**, 6176-6186.

Qin, S., Xing, K., Jiang, J-H., Xu, L-H., Li W-J. (2011) Biodiversity, bioactive natural products and biotechnological potential of plant-associated endophytic actinobacteria. *Appl. Microbiol. Biotechnol*. **89**, 457-473.

Raynaud, X. and Nunan, N. (2014) Spatial Ecology of Bacteria at the Microscale in Soil. *PLoS ONE* **9,** e87217. doi:10.1371/journal.pone.0087217.

Rebelo, A.G., Helme, N.A., Holmes, P.M., Forshaw, C.N., Richardson, S.H., Raimondo, D., Euston-Brown, D.I.W., Victor, J.E., Foden, W., Ebrahim, I., Bomhard, B., Oliver, E.G.H., Johns, A., van der Venter, J., van der Walt, R., von Witt, C., Low, A.B., Paterson-Jones, C., Rourke, J.P., Hitchcock, A.N., Potter, L., Vlok, J.H. & Pillay, D. (2005). *Leucospermum conocarpodendron* (L.) H.Buek subsp. *conocarpodendron*. National Assessment: Red List of South African Plants version 2014.1. Accessed on November 2014. http://redlist.sanbi.org/species.php?species=795-15.

Riesenfeld, C.S., Schloss, P.D., and Handelsman, J. (2004). Metagenomics: Genomic Analysis of Microbial Communities. *Annu. Rev. Genet*. **38**, 525-552.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). In *Molecular Cloning, a laboratory manual*, second edition. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Schmeisser, C., Steel, H. and Streit, W.R. (2007). Metagenomics, biotechnology with non-culturable microbes. *Appl. Microbiol. Biot*. **75**, 955-962

Schutte-Vlok, A.L. and Raimondo, D. (2012). *Podalyria sericea* (Andrews) R.Br. ex Aiton f. National Assessment: Red List of South African Plants version 2014.1. Accessed on November 2014. http://redlist.sanbi.org/species.php?species=446-46.

Senthikumar, M., Anandham, R., Madhaiyan, M. and Venkateswaran, V. (2011). Endophytic bacteria: Perspectives and applications in agricultural crop production.In Maheshwari, D.K. (Ed.) *Bacteria in AgroBiology: Crop Ecosystems*. New York, U.S.A. Springer.

Shenpagam, N.H., Kanchana, D.D., Sinduja G., Sandhya, R. (2012). Isolation of Endophytic Actinomycetes from Medicinal Plants and its Mutational effect in Biocontrol activity. *Int. J. Pharm. Sci. Res*. **3**, 4338-4344.

Shirling, E.B. and Gottlieb, D. (1966). Methods for characterization of *Streptomyces* species. *Int. J. Syst. Bacteriol*. **16**, 313-340.

Slanetz, L.W., Bent, D.F. and Bartley, C.H. (1955). Use of the Membrane Filter Technique to Enumerate Enterococci in Water. *Public Health Reports*. **7**, 67-72.

Sohn, K., Hong, S.G., Bae, K.S. and Chun, J. (2003). Transfer of *Hongia koreensis* Lee *et al*. 2000 to the genus *Kribbella* Park *et al.* 1999 as *Kribbella koreensis* comb.nov. *Int. J. Syst. Evol. Microbiol.* **53**, 1005-1007.

Strobel, G. (2003). Endophytes as sources of bioactive products. *Microbes Infect.* **5**, 535–544.

Tan, G.Y.A., Ward, A.C. and Goodfellow, M. (2006). Exploration of *Amycolatopsis* diversity in soil using genus-specific primers and novel selective media. *Syst. Appl. Microbiol*. **29**, 557-569.

Tanvir, R., Sajid, I., Hasnain, S. (2014). Biotechnological potential of endophytic actinomycetes associated with Asteraceae plants: isolation, biodiversity and bioactivities. *Biotechnol. Lett*. **36**, 767–773.

Taylor, J.E., Lee, S. and Crous, P.W. (2001). Biodiversity in the Cape Floral Kingdom: fungi occurring on *Proteaceae. Mycol. Res*. **105**, 1480-1484.

Trujillo, M.E., Kroppenstedt, R.M., Schumann, P., and Martinez-Molina, E. (2006). *Kribbella lupini* sp. nov., isolated from the roots of *Lupinus angustifolius*. *Int. J. Syst. Evol. Microbiol*. **56**, 407-411.

University of Connecticut (2014). Accession Data; *Protea aurea ssp. aurea* (Burm.f.) Rourke. Accessed May 2014. http://florawww.eeb.uconn.edu/200800189.html.

Urzí, C., De Leo, F. and Schumann, P. (2008). *Kribbella catacumbae* sp. nov. and *Kribbella sancticallisti* sp. nov., isolated from whitish-grey patinas in the catacombs of St. Callistus in Rome, Italy. *Int. J. Syst. Evol. Microbiol.* **58**, 2090-2097.

Van der Walt, L. (2000). *Orphium frutescens*. PlantZAfrica.com. South African National Biodiversity Institute. Accessed May 2014. http://www.plantzafrica.com/frames/plantsfram.htm

Viljoen, C. (2007). *Podalyria sericea*. PlantZAfrica.com. South African National Biodiversity Institute. Accessed May 2014. http://www.plantzafrica.com/frames/plantsfram.htm

Wang, Y., Zhang, Z. and Ruan, J. (1996). A Proposal to Transfer *Microbispora bispora* (Lechevalier 1965) to a New Genus, *Thermobispora* gen. nov., as *Thermobispora bispora* comb. nov. *Int. J. Syst. Bacteriol.* **46**, 933-938.

Wayne, L.G., Brenner, D.J., Colwell, R.R., Grimont, P.A.D., Kandler, O., Krichevsky, M.I., Moore, L.H., Moore, W.E.C., Murray, R.G.E., Stackebrandt, E., Starr, M.P. and Trüper, H.G. (1987). International Committee on Systematic Bacteriology. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *Int. J. Syst. Bacteriol*. **37**, 463-464.

2.7 Appendix

Appendix 2A: *gyrB* genetic distance values for the genus *Kribbella* based on the 245 bp section within the variable region

	Kribbella strains	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
1	**PY4**																															
2	**SH5**	0.000																														
3	**SY9**	0.000	0.000																													
4	***Kribbella alba* DSM15500^T**	0.115	0.115	0.115																												
5	***Kribbella aluminosa* DSM18824^T**	0.042	0.042	0.042	0.115																											
6	***Kribbella aluminosa* HKI0480**	0.042	0.042	0.042	0.115	0.000																										
7	***Kribbella aluminosa* HKI0479**	0.042	0.042	0.042	0.115	0.000	0.000																									
8	***Kribbella amoyensis* NRBC107914^T**	0.073	0.073	0.073	0.100	0.082	0.082	0.082																								
9	***Kribbella anitibiotica* DSM15501^T**	0.008	0.008	0.008	0.124	0.051	0.051	0.051	0.082																							
10	***Kribbella albertanoniae* BC640^T**	0.096	0.096	0.096	0.077	0.100	0.100	0.100	0.086	0.100																						
11	***Kribbella catacumbae* DSM19601^T**	0.105	0.105	0.105	0.073	0.129	0.129	0.129	0.114	0.110	0.051																					
12	***Kribbella catacumbae* BC628**	0.105	0.105	0.105	0.073	0.129	0.129	0.129	0.114	0.110	0.051	0.000																				
13	***Kribbella catacumbae* BC630**	0.105	0.105	0.105	0.073	0.129	0.129	0.129	0.114	0.110	0.051	0.000	0.000																			
14	***Kribbella catacumbae* BC632**	0.105	0.105	0.105	0.073	0.129	0.129	0.129	0.114	0.110	0.051	0.000	0.000	0.000																		
15	***Kribbella endophytica* DSM23718^T**	0.105	0.105	0.105	0.087	0.119	0.119	0.119	0.105	0.110	0.077	0.059	0.059	0.059	0.059																	
16	***Kribbella italica* BC637^T**	0.082	0.082	0.082	0.082	0.100	0.100	0.100	0.086	0.091	0.082	0.055	0.055	0.055	0.055	0.029																
17	***Kribbella flavida* CIP107494^T**	0.073	0.073	0.073	0.077	0.087	0.087	0.087	0.082	0.082	0.064	0.069	0.069	0.069	0.069	0.073	0.068															
18	***Kribbella ginsengisoli* DSM 17941^T**	0.110	0.110	0.110	0.051	0.124	0.124	0.124	0.114	0.115	0.068	0.042	0.042	0.042	0.042	0.082	0.082	0.064														
19	***Kribbella hippodromi* S1.4^T**	0.025	0.025	0.025	0.119	0.033	0.033	0.033	0.082	0.029	0.091	0.110	0.110	0.110	0.110	0.105	0.091	0.073	0.110													
20	***Kribbella jejuensis* CIP108509^T**	0.034	0.034	0.034	0.114	0.033	0.033	0.033	0.086	0.042	0.096	0.114	0.114	0.114	0.114	0.096	0.086	0.068	0.124	0.033												
21	***Kribbella karoonensis* Q41^T**	0.042	0.042	0.042	0.129	0.029	0.029	0.029	0.068	0.051	0.114	0.129	0.129	0.129	0.129	0.129	0.101	0.087	0.134	0.033	0.046											
22	***Kribbella koreensis* CIP108301^T**	0.110	0.110	0.110	0.046	0.119	0.119	0.119	0.110	0.115	0.064	0.042	0.042	0.042	0.042	0.077	0.082	0.068	0.004	0.110	0.119	0.134										
23	***Kribbella lupini* LU14^T**	0.100	0.100	0.100	0.082	0.110	0.110	0.110	0.105	0.110	0.082	0.055	0.055	0.055	0.055	0.029	0.017	0.051	0.082	0.101	0.087	0.110	0.082									
24	***Kribbella sancticallisti* DSM19602^T**	0.087	0.087	0.087	0.046	0.096	0.096	0.096	0.082	0.096	0.064	0.051	0.051	0.051	0.051	0.055	0.042	0.042	0.060	0.096	0.091	0.101	0.064	0.042								
25	***Kribbella sancticallisti* BC634**	0.087	0.087	0.087	0.046	0.096	0.096	0.096	0.082	0.096	0.064	0.051	0.051	0.051	0.051	0.055	0.042	0.042	0.060	0.096	0.091	0.101	0.064	0.042	0.000							
26	***Kribbella sandramycini* DSM15626^T**	0.086	0.086	0.086	0.082	0.077	0.077	0.077	0.110	0.096	0.119	0.100	0.100	0.100	0.100	0.110	0.082	0.110	0.100	0.077	0.073	0.077	0.100	0.091	0.091	0.091						
27	***Kribbella shirazensis* DSM45490^T**	0.033	0.033	0.033	0.114	0.033	0.033	0.033	0.068	0.042	0.096	0.114	0.114	0.114	0.114	0.105	0.091	0.077	0.114	0.033	0.034	0.029	0.110	0.100	0.100	0.100	0.073					
28	***Kribbella solani* CIP108508^T**	0.000	0.000	0.000	0.115	0.042	0.042	0.042	0.073	0.008	0.096	0.105	0.105	0.105	0.105	0.105	0.082	0.073	0.110	0.025	0.034	0.042	0.110	0.100	0.087	0.087	0.086	0.033				
29	***Kribbella solani* YB2**	0.000	0.000	0.000	0.115	0.042	0.042	0.042	0.073	0.008	0.096	0.105	0.105	0.105	0.105	0.105	0.082	0.073	0.110	0.025	0.034	0.042	0.110	0.100	0.087	0.087	0.086	0.033	0.000			
30	***Kribbella swartbergensis* HMC25^T**	0.055	0.055	0.055	0.110	0.055	0.055	0.055	0.059	0.064	0.091	0.115	0.115	0.115	0.115	0.115	0.096	0.082	0.129	0.064	0.051	0.059	0.124	0.105	0.096	0.096	0.073	0.046	0.055	0.055		
31	***Kribbella yunnanensis* DSM15499^T**	0.096	0.096	0.096	0.068	0.105	0.105	0.105	0.100	0.100	0.038	0.033	0.033	0.033	0.033	0.055	0.055	0.060	0.042	0.096	0.100	0.119	0.046	0.055	0.033	0.033	0.100	0.100	0.096	0.096	0.105	
32	***Microlunatus phosphovorus* NM-1^T**	0.227	0.227	0.227	0.211	0.233	0.233	0.233	0.233	0.233	0.200	0.227	0.227	0.227	0.227	0.244	0.244	0.227	0.200	0.211	0.227	0.222	0.205	0.267	0.238	0.238	0.211	0.222	0.227	0.227	0.233	0.205

Chapter 3:
Metagenomic screening for the genus *Kribbella* in soil and plant samples

Chapter 3: Metagenomic screening for the genus *Kribbella* in soil and plant samples

Chapter 3: Metagenomic screening for the genus *Kribbella* in soil and plant samples

3.1 Summary

To enable a targeted search to isolate *Kribbella* strains, a metagenomic method was developed to indicate the presence of *Kribbella* DNA in environmental samples. This culture-independent approach is fast and can be applied to a sample before plate isolation techniques take place. A *Kribbella*-specific 16S rRNA gene PCR primer (Krb977F) was designed and the primer combination Krb977F+R5 was used to amplify *Kribbella* 16S rRNA gene sequences directly from DNA extracted from soil and plant samples. Each amplified PCR product was cloned into a plasmid and transformed into *Escherichia coli* cells. The *E. coli* colonies were screened by PCR to ensure the correct sized insert was present. Transformants containing inserts were grown in broth culture and the plasmids were extracted and sequenced. The edited sequences were analyzed by BLAST against the GenBank and Ez-Taxon-e databases to determine the organism in the environmental sample from which the 16S rRNA gene sequence had been cloned. In this way, an environmental sample can be screened to determine if *Kribbella* 16S rRNA genes are present before selective isolation on agar plates is attempted. *Kribbella* 16S rRNA gene sequences were successfully amplified from two soil samples and each of five leaf and five root samples. The Krb977F primer proved to be highly specific for *Kribbella* sequences, as only nine sequences belonging to cultured type strains of other genera were amplified. Of a total of 330 metagenomic 16S rRNA gene sequences, 181 (54.8%) were identified as belonging to *Kribbella*, 9 (2.7%) were identified as non-*Kribbella* cultured strains and 140 (42.4%) were most closely related to uncultured strains. The data also showed that *Kribbella* sequences were detected in both leaf and root samples of all the sampled indigenous plants (*Agathosma ovata* 'Glentana', *Leucospermum conocarpodendron*, *Orphium frutescens, Podylaria sericea* and *Protea aurea* subspecies *aurea*).

3.2 Introduction

It is well accepted that most microbes in nature have not been studied or cultured under laboratory conditions [Riesenfeld *et al.,* 2004; Rondon *et al.,* 2000; Schmeisser *et al.,* 2007]. It is estimated that culturable microbes account for only 1% or less of the bacterial diversity in most environmental samples and that this hidden diversity represents an intriguing resource for the development of novel genes, enzymes, antibiotics and other secondary metabolites for use in biotechnology [Riesenfeld *et al.,* 2004; Schmeisser *et al.,* 2007]. Metagenomics is the tool by which the total genetic pool of all the microbes in a particular environment may be assessed and investigated in a culture-independent manner [Culligan *et al*., 2014; Schmeisser *et al*., 2007]. The term is derived from the statistical concept of meta-analysis and genomics and involves cloning large fragments of DNA isolated directly from microbes in natural environments [Rondon *et al*., 2000; Schloss and Handelsman, 2003; Schmeisser *et al.,* 2007]. It has been used successfully as a tool on a variety of levels, from the study of single genes, pathways, organisms and entire communities [Riesenfeld *et al*., 2004]. Metagenomics has been invaluable in the study of the human gut microbiome and has lead to greater understanding of changes in the gut's microbial communities under various circumstances [Graessler *et al.,* 2013; Lepage *et al*., 2013; Levy and Borenstein, 2014; Raes, 2014; Turnbaugh *et al.,* 2009]. In this study, metagenomic analysis was used to determine the presence of a single actinobacterial genus living in soil and in various plant tissues.

Actinobacteria that live in the tissues of living plants and do not harm the plants are known as endophytic actinobacteria [Qin *et al*., 2009]. Relatively few endophytic actinobacteria have been studied, indicating a great opportunity to find novel species and novel secondary metabolites among different plants and ecosystems [Qin *et al.,* 2011]. As a result of the high rate of rediscovery of secondary metabolites, isolating rarer or lesser-known actinomycetes present in this underexplored plant niche is of significant interest in the search for novel compounds [Kurtböke, 2012; Qin *et al.,* 2011]. These rare actinomycetes can produce diverse and complex compounds: from 58 rare actinomycete genera, over 2500 bioactive compounds have been isolated [Bérdy, 2005;

Kurtböke, 2012]. A better understanding of the ecology of these actinomycetes, their effective isolation from environmental sources and their rapid and reliable identification are needed in order to fully utilize and benefit from them [Kurtböke, 2012]. As a rare genus, *Kribbella* is an ideal candidate to serve as an example in the use of metagenomics as a tool to screen environmental samples. The novel endophytic taxon, '*Kribbella podocarpi*' strain YPL1, from the South African tree *Podocarpus latifolius* and the newly described species *Kribbella endophytica*, (isolated from a native Australian apricot tree, *Pittosporum angustifolium)* provide evidence that *Kribbella* species are living as endophytes in a variety of plants and that metagenomic screening for the presence of *Kribbella* may be employed alongside plate isolation techniques [Kaewkla and Franco, 2013; Curtis *et al*., manuscript under revision].

Ecosystems with the greatest general biodiversity also seem to have the greatest diversity of endophytes. This provides strong motivation to search for endophytic actinobacteria in plants belonging to the highly biodiverse Cape Floral Kingdom [Strobel *et al*., 2004]. It was the aim of this part of the study to search this underexplored reservoir for the presence of *Kribbella* DNA by screening the five indigenous fynbos plants that were used to isolate *Kribbella* colonies (described in Chapter 2).

3.3 Methods and Materials

3.3.1 Plant samples

The plant samples consisted of surface sterilized roots and leaves from the following plants: *Orphium frutescens, Leucospermum conocarpodendron*, *Podylaria sericea*, *Protea aurea* subspecies *aurea* and *Agathosma ovata* 'Glentana'. These were the same plants used for the culture-dependent isolation of *Kribbella* colonies described in Chapter 2.

For *A. ovata*, in addition to the root and leaf samples, a sample was also taken from the soil around the plant for metagenomic screening. A second soil sample, from the Tokai forest (Tokai suburb, Cape Town), was also used to try to isolate *Kribbella* colonies (on

ISP2, 7H9, CZ and modified SM1 plates) and for metagenomic screening for the presence of *Kribbella*.

3.3.2 16S rRNA gene primer design

The *Kribbella*-specific primer was designed from an alignment generated in MEGA 5.05 [Tamura *et al.,* 2011] of the 16S rRNA gene sequences of the 16 *Kribbella* type strains' that were available at the time (March 2011). The *Kribbella* sequences were used to define a consensus 16S rRNA gene sequence for the genus. This consensus sequence was then compared to 16S rRNA gene sequences belonging to members of the families *Nocardioidaceae* and *Propionibacteriaceae* in a second alignment (the sequences used are listed in Appendix 3A). This second alignment was then analyzed manually for nucleotide sequences unique to the *Kribbella* consensus sequence and two 16S rRNA gene primers were designed. The forward primer Krb977F (5'-TAC CTG GGT TTG ACA TAT AGG GAA ATC-3') binds to positions 977-1003 of the 16S rRNA gene of *Kribbella flavida* DSM 17836^T (CP001736) and is used in combination with the R5 universal 16S rRNA gene primer [Cook and Meyers, 2003]. The reverse primer Krb624R (5'-AGT TCC GCT CTC CCC TGC CTA C-3') binds to positions 645-624 of the *K. flavida* DSM 17836^T (CP001736) 16S rRNA gene and is used with the universal F1 16S rRNA gene primer [Cook and Meyers, 2003]. The binding positions of the new primers on the 16S rRNA gene are such that they cannot be used together.

The new primers (Krb977F & R5 and F1 & Krb624R) were tested on all *Kribbella* type strains to ensure amplification of fragments of the expected sizes. The primers were also tested against various other genera (namely, *Herbidospora, Nonomuraea, Planomonospora, Sphaerisporangium* and *Streptosporangium* from the family *Streptosporangiaceae*), to ensure specific binding to only the *Kribbella* genus (no strains from genera other than *Kribbella* in the family *Nocardioidaceae* were available for testing). The reverse primer amplified 16S rRNA gene sequences of all genera mentioned (as can be seen in Appendix 3B) and was abandoned as it was not specific for *Kribbella*. The Krb977F/R5 primer combination proved to amplify only *Kribbella* 16S rRNA gene sequences, producing a 535-bp fragment.

3.3.3 Sample DNA extraction

Soil, leaf and root samples were used as the starting environmental samples and are described in Chapter 2. Total DNA was extracted directly from 0.5g of the soil samples with the Zymo Research Soil Microbe DNA Mini Prep kit (Irvine, USA) following the manufacturer's protocol. The leaf and root material of each plant was first surface sterilized following the procedure described in Chapter 2 and then the leaves and roots were cut into < 5mm wide pieces and incubated in quarter strength phosphate buffer solution for 30 minutes with shaking at 30°C [Kirby and Meyers, 2010]. Each root and each leaf sample was incubated separately in quarter strength phosphate buffer. DNA was extracted from 500 µl of the phosphate buffer suspension using the Zymo Research Soil Microbe DNA Mini Prep kit (Irvine, USA).

3.3.4 PCR amplifications of the 16S rRNA gene

The extracted soil, leaf or root DNA was then amplified by PCR. Initially for the soil samples, the universal 16S rRNA gene primers F1/R5 [Cook and Meyers, 2003] were used in a standard PCR program to amplify all 16S rRNA gene sequences present. A second, touchdown PCR program using the Krb977F/R5 primer combination was then used to amplify *Kribbella* specific sequences from the products amplified in the first round of PCR. In later experiments with the plant samples, it was found that an initial PCR amplification using the Krb977F/R5 primer combination (instead of the F1/R5 primer combination) resulted in clearer DNA bands. Therefore, the plant samples were initially amplified with a standard PCR program using the Krb977F/R5 primers and then the same primers were used again in the touchdown PCR program.

All standard PCR reactions were carried out in 50µl volumes. PCR experiments were initially run using SuperTherm *Taq* polymerase (JMR Holdings, USA), but later a switch was made to Bioline MyTaq *Taq* polymerase (Bioline London, United Kingdom). For SuperTherm PCR, each reaction contained 2 mM $MgCl_2$, 0.5U SuperTherm *Taq* polymerase, 150 µM of each dNTP, 0.5 µM of each primer and 500ng template DNA. For Bioline MyTaq PCR, 1U of *Taq* polymerase was used in each 50µl reaction with 3mM $MgCl_2$, 250µM of each dNTP, 0.5 µM of each primer and 500ng template DNA. Colony PCR reactions were carried out in 20µl volumes.

PCR was performed using a Techne TC-512 thermal cycler. For the plant samples, the Krb977F/R5 primers were used in both sets of PCR programs. SuperTherm *Taq* polymerase was used for the Tokai forest soil sample. The standard PCR program using SuperTherm *Taq* was: an initial denaturation of 96°C for 2 min, 30 cycles of denaturation at 96°C for 45s, annealing at 56°C for 30s, extension at 72°C for 1 minute 30s and a final extension at 72°C for 5 minutes. The standard PCR program using MyTaq *Taq* polymerase was: an initial denaturation at 95°C for 1 minute 40 seconds, 30 cycles of denaturation at 95°C for 15s, annealing at 55°C for 15s, extension at 72°C for 10s and a final extension at 72°C for 5 minutes. The PCR touchdown program with SuperTherm *Taq* was: an initial denaturation at 96°C for 2 minutes, 10 cycles of denaturation at 95°C for 45s, annealing at 62-60°C for 45s and extension at 72°C for 1 minute, followed by 15 cycles of denaturation at 95°C for 45s, annealing at 56°C for 45s and extension at 72°C for 1 minute followed by a final extension at 72°C for 5 minutes. The PCR touchdown program with MyTaq *Taq* was identical to the SuperTherm protocol except that the initial denaturation was at 95°C for 1 minute 40 seconds, the denaturation and annealing steps were each 15s long and the extension step was 10 seconds long. All PCR products were electrophoresed on 0.8% (w/v) agarose gels containing 0.8µg/ml ethidium bromide. The agarose gels were visualized on a GelDoc XR System (BioRad).

3.3.5 Cloning, sequencing and sequence analysis

Each PCR product was purified using an MSB® Spin PCRapace kit (Invitek, Berlin, Germany) and ligated into the Promega pGEM-T Easy vector system and transformed into *E. coli* DH5alpha cells (Bioline, London, United Kingdom). The transformed cells were then plated on Luria-Bertani agar (LA) plates [Gerhardt *et al.,* 1994; Lennox, 1955] containing 80µg/ml X-gal, 0.5mM IPTG and 100µg/ml of ampicillin for blue/white colony selection. White colonies were selected and sub-cultured onto fresh LA plates containing 100µg/ml of ampicillin. The colonies were screened for the correct sized insert by colony PCR with the primers Krb977F/R5. The colonies that contained an insert of the correct size (535 bp) were then grown in Luria-Bertani broth containing 100µg/ml of ampicillin. The plasmids were extracted using the Zymo Research Zyppy™ Plasmid Miniprep Kit (Irvine, USA) following the manufacturer's protocol and sent for sequencing using the M13 sequencing primers [Messing, 1983]. Sequencing was performed as a service by Macrogen Inc., Seoul, South Korea. The sequence

chromatograms were edited manually with Chromas version 2.01 (Technelysium Pty. Ltd., Australia) to remove vector sequences and analyzed by BLAST against the GenBank and Ez-Taxon-e databases (both of which include sequences from cultured and uncultured strains) to determine the most closely related strains to the cloned sequence [Kim *et. al*., 2012].

3.4 Results

The recombinant *E. coli* colonies, totalling 815, were screened with the primers Krb977F/R5 to check for an insert of the appropriate size (535 bp) and those with the correct sized insert were sent for sequencing. A total of 330 clone sequences was obtained: *L. conocarpodendron*: 87 clone sequences, *A. ovata* var. 'Glentana': 68 clone sequences, *O. frutescens*: 71 clone sequences: *P. aurea* subspecies *aurea*: 39 clone sequences and *P. sericea*: 20 clone sequences; 45 clone sequences were obtained from the two soil samples. The identification of each clone sequence was accomplished using the Ez-Taxon-e database and the GenBank database and the results are listed in Appendix 3C. The top clone-sequence hits from the Ez-Taxon-e database can be seen in Table 3.1 [Kim *et. al*., 2012]. Only the 190 (57.6%) clone sequences identified as being closely related to a cultured type strain are displayed; all hits to uncultured bacteria have been removed. All sequences identified as *Kribbella* are shown in blue, with sequences having a 100% sequence similarity highlighted in green. Any *Kribbella* sequences with a low (<98%) sequence similarity are highlighted in orange.

The Krb977F primer proved successful in amplifying *Kribbella* 16S rRNA genes from all environmental samples. A large percentage (42.4%) of the clone sequences are most closely related to uncultured strains, which is not uncommon in metagenomic studies, as most microbes have not been cultured [Reisenfeld *et al.,* 2004; Schloss and Handelsman, 2003]. The Krb977F primer also proved to be highly specific for *Kribbella* sequences, as can be seen from Table 3.1, as only nine of 190 sequences (4.7%) were identified as being from non-*Kribbella* cultured strains: six from the plant samples and three from the soil samples.

From Table 3.1, several different *Kribbella* species were detected, with nine of the 20 type strains being represented within the 181 sequences identified as belonging to the genus *Kribbella*. A total of 15% of the *Kribbella* sequences shared a 100% sequence similarity with one or more *Kribbella* type strains namely, *Kribbella aluminosa, Kribbella karoonensis, Kribbella hippodromi* and *Kribbella solani*, which reveals a prevalence of these strains in the samples studied. The majority of the clone sequence similarities were under 100%. The lower sequence similarity of some of the clone sequences to cultured type strains (highlighted in orange) may indicate possible new species present within the samples. However, it must be borne in mind that short 16S rRNA gene sequences were used in this analysis.

Table 3.1. Ez-Taxon-e sequence similarity of metagenomic 16S rRNA gene clones to cultured strains.

Plant Name	Plant material	Clone Name	EZ-Taxon -e Top Hit	% Similarity
Podylaria sericea	Leaf	PodsL16	*Ilumatobacter fluminis* YM22-133(T)	96.6
Podylaria sericea	Leaf	PodsL25	***Kribbella hippodromi* S1.4(T)**	98.6
Podylaria sericea	Leaf	PodsL34	***Kribbella sancticallisti* BC633(T)**	99.2
Podylaria sericea	Root	PodsR5	***Kribbella aluminosa* HKI 0478(T)**	99.8
Podylaria sericea	Root	PodsR6	***Kribbella hippodromi* S1.4(T)**	99.0
Podylaria sericea	Root	PodsR8	***Kribbella hippodromi* S1.4(T)**	99.0
Podylaria sericea	Root	PodsR43	***Kribbella solani* DSA1(T)**	98.3
Podylaria sericea	Root	PodsR45	***Kribbella karoonensis* DSM 17344(T)**	97.3
Podylaria sericea	Root	PodsR46	***Kribbella hippodromi* S1.4(T)**	99.2
Podylaria sericea	Root	PodsR49	***Kribbella solani* DSA1(T)**	98.1
Podylaria sericea	Root	PodsR50	***Kribbella karoonensis* DSM 17344(T)**	98.5
Podylaria sericea	Root	PodsR58	***Kribbella hippodromi* S1.4(T)**	99.6
Podylaria sericea	Root	PodsR59	***Kribbella karoonensis* DSM 17344(T)**	97.3
Podylaria sericea	Root	PodsR60	***Kribbella hippodromi* S1.4(T)**	99.4
Podylaria sericea	Root	PodsR62	***Kribbella karoonensis* DSM 17344(T)**	95.3
Podylaria sericea	Root	PodsR73	***Kribbella hippodromi* S1.4(T)**	99.2
Protea aurea subsp aurea	Leaf	ProtaL23	***Kribbella sancticallisti* BC633(T)**	99.0
Protea aurea subsp aurea	Leaf	ProtaL32	***Kribbella sancticallisti* BC633(T)**	99.2
Protea aurea subsp aurea	Leaf	ProtaL42	***Kribbella hippodromi* S1.4(T)**	99.6
Protea aurea subsp aurea	Leaf	ProtaL49	***Kribbella sancticallisti* BC633(T)**	99.2
Protea aurea subsp aurea	Leaf	ProtaL71	***Kribbella karoonensis* DSM 17344(T)**	99.5
Protea aurea subsp aurea	Root	ProtaR42	***Kribbella hippodromi* S1.4(T)**	99.6
Protea aurea subsp aurea	Root	ProtaR65	*Ilumatobacter fluminis* YM22-133(T)	96.8
Leucospermum conocarpodendron	Leaf	LcL1	***Kribbella karoonensis* DSM 17344(T)**	92.4
Leucospermum conocarpodendron	Leaf	LcL2	***Kribbella solani* DSA1(T)**	99.6
Leucospermum conocarpodendron	Leaf	LcL4	***Kribbella aluminosa* HKI 0478(T)**	99.8
Leucospermum conocarpodendron	Leaf	LcL5	***Kribbella aluminosa* HKI 0478(T)**	99.8
Leucospermum conocarpodendron	Leaf	LcL6	***Kribbella aluminosa* HKI 0478(T)**	99.4
Leucospermum conocarpodendron	Leaf	LcL11	***Kribbella aluminosa* HKI 0478(T)**	99.8
Leucospermum conocarpodendron	Leaf	LcL12	***Kribbella sancticallisti* BC633(T)**	99.2
Leucospermum conocarpodendron	Leaf	LcL13	***Kribbella sancticallisti* BC633(T)**	99.2
Leucospermum conocarpodendron	Leaf	LcL15	***Kribbella solani* DSA1(T)**	99.6
Leucospermum conocarpodendron	Leaf	LcL16	***Kribbella karoonensis* DSM 17344(T)**	99.8
Leucospermum conocarpodendron	Leaf	LcL17	***Kribbella swartbergensis* HMC25(T)**	98.9
Leucospermum conocarpodendron	Leaf	LcL20	***Kribbella karoonensis* DSM 17344(T)**	93.0
Leucospermum conocarpodendron	Leaf	LcL21	***Kribbella sancticallisti* BC633(T)**	99.4
Leucospermum conocarpodendron	Leaf	LcL23	***Kribbella aluminosa* HKI 0478(T)**	99.6
Leucospermum conocarpodendron	Leaf	LcL25	***Kribbella aluminosa* HKI 0478(T)**	99.8
Leucospermum conocarpodendron	Leaf	LcL30	***Kribbella sancticallisti* BC633(T)**	99.0
Leucospermum conocarpodendron	Leaf	LcL31	***Kribbella karoonensis* DSM 17344(T)**	99.5
Leucospermum conocarpodendron	Leaf	LcL33	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	**100.0**
Leucospermum conocarpodendron	Leaf	LcL38	***Kribbella karoonensis* DSM 17344(T)**	99.8
Leucospermum conocarpodendron	Leaf	LcL39	***Kribbella karoonensis* DSM 17344(T)**	99.5

Table 3.1 continued

Plant Name	Plant material	Clone Name	EZ-Taxon -e Top Hit	% Similarity
Leucospermum conocarpodendron	Leaf	LcL41	***Kribbella shirazensis* UTMC 693(T)**	81.4
Leucospermum conocarpodendron	Leaf	LcL42	***Kribbella karoonensis* DSM 17344(T)**	91.1
Leucospermum conocarpodendron	Leaf	LcL44	***Kribbella aluminosa* HKI 0478(T)**	99.8
Leucospermum conocarpodendron	Leaf	LcL45	***Kribbella solani* DSA1(T)**	99.6
Leucospermum conocarpodendron	Leaf	LcL47	***Kribbella sancticallisti* BC633(T)**	99.4
Leucospermum conocarpodendron	Leaf	LcL48	***Kribbella sancticallisti* BC633(T)**	99.2
Leucospermum conocarpodendron	Leaf	LcL49	***Kribbella aluminosa* HKI 0478(T)**	99.8
Leucospermum conocarpodendron	Leaf	LcL50	***Kribbella solani* DSA1(T)**	99.6
Leucospermum conocarpodendron	Leaf	LcL53	***Kribbella aluminosa* HKI 0478(T)**	99.1
Leucospermum conocarpodendron	Leaf	LcL56	***Kribbella aluminosa* HKI 0478(T)**	99.6
Leucospermum conocarpodendron	Leaf	LcL57	***Kribbella sancticallisti* BC633(T)**	99.4
Leucospermum conocarpodendron	Leaf	LcL60	***Kribbella aluminosa* HKI 0478(T)**	**100.0**
Leucospermum conocarpodendron	Leaf	LcL71	***Kribbella sancticallisti* BC633(T)**	99.2
Leucospermum conocarpodendron	Root	LcR1	***Kribbella aluminosa* HKI 0478(T)**	98.7
Leucospermum conocarpodendron	Root	LcR3	***Kribbella solani* DSA1(T)**	99.8
Leucospermum conocarpodendron	Root	LcR8	***Kribbella hippodromi* S1.4(T)**	**100.0**
Leucospermum conocarpodendron	Root	LcR11	***Kribbella aluminosa* HKI 0478(T)**	99.6
Leucospermum conocarpodendron	Root	LcR12	***Kribbella aluminosa* HKI 0478(T)**	**100.0**
Leucospermum conocarpodendron	Root	LcR14	***Kribbella aluminosa* HKI 0478(T)**	99.8
Leucospermum conocarpodendron	Root	LcR15	***Kribbella aluminosa* HKI 0478(T)**	99.6
Leucospermum conocarpodendron	Root	LcR19	***Kribbella aluminosa* HKI 0478(T)**	99.8
Leucospermum conocarpodendron	Root	LcR20	***Kribbella aluminosa* HKI 0478(T)**	99.6
Leucospermum conocarpodendron	Root	LcR22	***Kribbella aluminosa* HKI 0478(T)**	99.6
Leucospermum conocarpodendron	Root	LcR23	***Kribbella aluminosa* HKI 0478(T)**	99.8
Leucospermum conocarpodendron	Root	LcR24	***Kribbella solani* DSA1(T)**	98.3
Leucospermum conocarpodendron	Root	LcR25	***Kribbella hippodromi* S1.4(T)**	**100.0**
Leucospermum conocarpodendron	Root	LcR28	***Kribbella aluminosa* HKI 0478(T)**	**100.0**
Leucospermum conocarpodendron	Root	LcR34	***Kribbella karoonensis* DSM 17344(T)**	99.5
Leucospermum conocarpodendron	Root	LcR35	***Kribbella sancticallisti* BC633(T)**	99.2
Leucospermum conocarpodendron	Root	LcR40	***Kribbella aluminosa* HKI 0478(T) / *Kribbella karoonensis* DSM 17344(T)**	**100.0**
Leucospermum conocarpodendron	Root	LcR42	***Kribbella sancticallisti* BC633(T)**	99.0
Leucospermum conocarpodendron	Root	LcR45	***Kribbella aluminosa* HKI 0478(T)**	**100.0**
Leucospermum conocarpodendron	Root	LcR49	***Kribbella swartbergensis* DSM 17345(T)**	98.8
Leucospermum conocarpodendron	Root	LcR50	***Kribbella aluminosa* HKI 0478(T) / *Kribbella karoonensis* DSM 17344(T)**	**100.0**
Leucospermum conocarpodendron	Root	LcR59	***Kribbella sancticallisti* BC633(T)**	98.4
Leucospermum conocarpodendron	Root	LcR60	***Kribbella aluminosa* HKI 0478(T)**	99.8
Leucospermum conocarpodendron	Root	LcR62	***Kribbella aluminosa* HKI 0478(T)**	99.8
Leucospermum conocarpodendron	Root	LcR63	***Kribbella amoyensis* XMU 198(T)**	99.4
Leucospermum conocarpodendron	Root	LcR65	***Kribbella sancticallisti* BC633(T)**	99.2
Orphium frutescens	Leaf	OfL2	***Kribbella solani* DSA1(T)**	99.1
Orphium frutescens	Leaf	OfL3	***Kribbella solani* DSA1(T) / *Kribbella aluminosa* HKI 0478(T)**	99.3
Orphium frutescens	Leaf	OfL5	***Kribbella shirazensis* UTMC 693(T)**	99.2
Orphium frutescens	Leaf	OfL7	***Kribbella karoonensis* DSM 17344(T)**	82.5
Orphium frutescens	Leaf	OfL8	***Kribbella aluminosa* HKI 0478(T)**	99.8
Orphium frutescens	Leaf	OfL12	***Kribbella solani* DSA1(T)**	94.5
Orphium frutescens	Leaf	OfL21	***Kribbella aluminosa* HKI 0478(T)**	99.8

Table 3.1 continued

Plant Name	Plant material	Clone Name	EZ-Taxon -e Top Hit	% Similarity
Orphium frutescens	Leaf	OfL31	***Kribbella aluminosa* HKI 0478(T)**	99.6
Orphium frutescens	Leaf	OfL64	***Kribbella aluminosa* HKI 0478(T)**	99.8
Orphium frutescens	Leaf	OfL65	***Kribbella sancticallisti* BC633(T)**	99.4
Orphium frutescens	Leaf	OfL68	***Kribbella sancticallisti* BC633(T)**	99.4
Orphium frutescens	Leaf	OfL69	***Kribbella sancticallisti* BC633(T)**	99.0
Orphium frutescens	Leaf	OfL72	***Kribbella aluminosa* HKI 0478(T)**	99.6
Orphium frutescens	Leaf	OfL73	***Kribbella aluminosa* HKI 0478(T)**	99.4
Orphium frutescens	Leaf	OfL75	***Kribbella sancticallisti* BC633(T)**	97.7
Orphium frutescens	Leaf	OfL78	***Kribbella solani* DSA1(T)**	98.7
Orphium frutescens	Leaf	OfL79	***Kribbella hippodromi* S1.4(T)**	99.4
Orphium frutescens	Root	OfR17	***Kribbella sandramycini* ATCC 39419(T)**	93.5
Orphium frutescens	Root	OfR21	***Kribbella swartbergensis* HMC25(T)**	99.4
Orphium frutescens	Root	OfR24	*Blastococcus aggregatus* ATCC 25902(T)	97.0
Orphium frutescens	Root	OfR30	***Kribbella sancticallisti* BC633(T)**	99.2
Orphium frutescens	Root	OfR31	***Kribbella sancticallisti* BC633(T)**	99.2
Orphium frutescens	Root	OfR33	***Kribbella aluminosa* HKI 0478(T)**	99.6
Orphium frutescens	Root	OfR35	***Kribbella sancticallisti* BC633(T)**	99.4
Orphium frutescens	Root	OfR36	***Kribbella aluminosa* HKI 0478(T) / *Kribbella karoonensis* DSM 17344(T)**	100.0
Orphium frutescens	Root	OfR39	***Kribbella solani* DSA1(T)**	99.6
Orphium frutescens	Root	OfR43	***Kribbella solani* DSA1(T)**	99.6
Orphium frutescens	Root	OfR45	***Kribbella sancticallisti* BC633(T)**	99.4
Orphium frutescens	Root	OfR63	***Kribbella sancticallisti* BC633(T)**	98.0
Orphium frutescens	Root	OfR64	***Kribbella aluminosa* HKI 0478(T) / *Kribbella karoonensis* DSM 17344(T)**	100.0
Orphium frutescens	Root	OfR65	***Kribbella aluminosa* HKI 0478(T)**	99.8
Orphium frutescens	Root	OfR66	***Kribbella aluminosa* HKI 0478(T)**	99.1
Orphium frutescens	Root	OfR69	***Kribbella solani* DSA1(T)**	99.4
Orphium frutescens	Root	OfR70	***Kribbella sancticallisti* BC633(T)**	99.2
Orphium frutescens	Root	OfR78	***Kribbella solani* DSA1(T)**	96.8
Agathosma ovata 'Glentana'	Leaf	AgL2	***Kribbella solani* DSA1(T)**	99.6
Agathosma ovata 'Glentana'	Leaf	AgL3	***Kribbella solani* DSA1(T)**	99.8
Agathosma ovata 'Glentana'	Leaf	AgL6	***Kribbella hippodromi* S1.4(T)**	99.8
Agathosma ovata 'Glentana'	Leaf	AgL12	***Kribbella solani* DSA1(T)**	99.8
Agathosma ovata 'Glentana'	Leaf	AgL13	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	100.0
Agathosma ovata 'Glentana'	Leaf	AgL 15	***Kribbella solani* DSA1(T) / *Kribbella aluminosa* HKI 0478(T)**	99.7
Agathosma ovata 'Glentana'	Leaf	AgL18	***Kribbella solani* DSA1(T)**	99.6
Agathosma ovata 'Glentana'	Leaf	AgL26	***Kribbella solani* DSA1(T)**	99.6
Agathosma ovata 'Glentana'	Leaf	AgL27	*Blastococcus aggregatus* ATCC 25902(T)	98.4
Agathosma ovata 'Glentana'	Leaf	AgL28	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	100.0
Agathosma ovata 'Glentana'	Leaf	AgL29	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	100.0
Agathosma ovata 'Glentana'	Leaf	AgL32	*Pseudomonas toyotomiensis* HT-3(T)	96.2
Agathosma ovata 'Glentana'	Leaf	AgL33	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	100.0
Agathosma ovata 'Glentana'	Leaf	AgL34	***Kribbella solani* DSA1(T)**	99.6
Agathosma ovata 'Glentana'	Leaf	AgL35	***Kribbella solani* DSA1(T)**	99.6
Agathosma ovata 'Glentana'	Leaf	AgL37	***Kribbella solani* DSA1(T)**	99.6
Agathosma ovata 'Glentana'	Leaf	AgL40	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	100.0
Agathosma ovata 'Glentana'	Leaf	AgL41	***Kribbella solani* DSA1(T)**	99.8

Table 3.1 continued

Plant Name	Plant material	Clone Name	EZ-Taxon -e Top Hit	% Similarity
Agathosma ovata 'Glentana'	Leaf	AgL42	***Kribbella solani* DSA1(T)**	99.6
Agathosma ovata 'Glentana'	Leaf	AgL44	***Kribbella solani* DSA1(T)**	99.6
Agathosma ovata 'Glentana'	Leaf	AgL51	***Kribbella solani* DSA1(T)**	99.6
Agathosma ovata 'Glentana'	Leaf	AgL52	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	**100.0**
Agathosma ovata 'Glentana'	Leaf	AgL53	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	**100.0**
Agathosma ovata 'Glentana'	Leaf	AgL54	***Kribbella solani* DSA1(T)**	99.4
Agathosma ovata 'Glentana'	Leaf	AgL55	***Kribbella karoonensis* DSM 17344(T)**	99.5
Agathosma ovata 'Glentana'	Leaf	AgL56	***Kribbella solani* DSA1(T)**	99.6
Agathosma ovata 'Glentana'	Leaf	AgL57	***Kribbella solani* DSA1(T)**	99.8
Agathosma ovata 'Glentana'	Leaf	AgL58	***Kribbella solani* DSA1(T)**	99.1
Agathosma ovata 'Glentana'	Leaf	AgL60	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	**100.0**
Agathosma ovata 'Glentana'	Leaf	AgL62	***Kribbella hippodromi* S1.4(T)**	**100.0**
Agathosma ovata 'Glentana'	Leaf	AgL63	***Kribbella solani* DSA1(T)**	99.8
Agathosma ovata 'Glentana'	Leaf	AgL64	***Kribbella solani* DSA1(T)**	99.3
Agathosma ovata 'Glentana'	Leaf	AgL65	***Kribbella solani* DSA1(T)**	99.4
Agathosma ovata 'Glentana'	Leaf	AgL66	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	**100.0**
Agathosma ovata 'Glentana'	Leaf	AgL73	***Kribbella solani* DSA1(T)**	99.4
Agathosma ovata 'Glentana'	Leaf	AgL74	***Kribbella solani* DSA1(T)**	99.8
Agathosma ovata 'Glentana'	Leaf	AgL75	***Kribbella solani* DSA1(T)**	99.4
Agathosma ovata 'Glentana'	Leaf	AgL76	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	**100.0**
Agathosma ovata 'Glentana'	Leaf	AgL77	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	**100.0**
Agathosma ovata 'Glentana'	Leaf	AgL79	***Kribbella solani* DSA1(T)**	99.4
Agathosma ovata 'Glentana'	Leaf	AgL80	***Kribbella solani* DSA1(T)**	99.8
Agathosma ovata 'Glentana'	Leaf	AgL82	***Kribbella solani* DSA1(T)**	99.4
Agathosma ovata 'Glentana'	Leaf	AgL83	***Kribbella solani* DSA1(T)**	99.8
Agathosma ovata 'Glentana'	Leaf	AgL84	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	**100.0**
Agathosma ovata 'Glentana'	Leaf	AgL85	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	**100.0**
Agathosma ovata 'Glentana'	Leaf	AgL87	***Kribbella solani* DSA1(T)**	99.6
Agathosma ovata 'Glentana'	Stem	AgS7	***Kribbella solani* DSA1(T)**	99.6
Agathosma ovata 'Glentana'	Stem	AgS9	*Blastococcus saxobsidens* BC448(T)	98.1
Agathosma ovata 'Glentana'	Root	AgRK1	***Kribbella karoonensis* DSM 17344(T)**	**100.0**
Agathosma ovata 'Glentana'	Root	AgRK7	***Kribbella aluminosa* HKI 0478(T)**	99.8
Agathosma ovata 'Glentana'	Root	AgRK11	***Kribbella aluminosa* HKI 0478(T) / *Kribbella karoonensis* DSM 17344(T)**	**100.0**
Agathosma ovata 'Glentana'	Root	AgRK13	***Kribbella karoonensis* DSM 17344(T)**	**100.0**
Agathosma ovata 'Glentana'	Root	AgRK14	***Kribbella karoonensis* DSM 17344(T)**	**100.0**
Agathosma ovata 'Glentana'	Root	AgRK17	***Kribbella aluminosa* HKI 0478(T) / *Kribbella karoonensis* DSM 17344(T)**	**100.0**
Agathosma ovata 'Glentana'	Root	AgRK24	***Kribbella aluminosa* HKI 0478(T)**	99.6
Agathosma ovata 'Glentana'	Root	AgRK26	***Kribbella aluminosa* HKI 0478(T)**	99.6
Agathosma ovata 'Glentana'	Root	AgRK30	***Kribbella sancticallisti* BC633(T)**	99.0
Agathosma ovata 'Glentana'	Root	AgRK31	***Kribbella shirazensis* UTMC 693(T)**	99.4
Agathosma ovata 'Glentana'	Root	AgRK39	***Kribbella sancticallisti* BC633(T)**	99.2
Agathosma ovata 'Glentana'	Root	AgRK40	***Kribbella karoonensis* DSM 17344(T)**	**100.0**
Agathosma ovata 'Glentana'	Root	AgRK49	***Kribbella sancticallisti* BC633(T)**	99.4
Agathosma ovata 'Glentana'	Root	AgRK50	***Kribbella aluminosa* HKI 0478(T) / *Kribbella karoonensis* DSM 17344(T)**	**100.0**
Agathosma ovata 'Glentana'	Root	AgRK53	***Kribbella shirazensis* UTMC 693(T)**	99.4
Agathosma ovata 'Glentana'	Root	AgRK67	***Kribbella shirazensis* UTMC 693(T)**	99.4

Table 3.1 continued

Plant Name	Plant material	Clone Name	EZ-Taxon -e Top Hit	% Similarity
Agathosma ovata 'Glentana'	Root	AgRK78	***Kribbella aluminosa* HKI 0478(T)**	99.7
Agathosma ovata 'Glentana'	Soil	K5 Ag S	***Kribbella aluminosa* HKI 0478(T)**	99.6
Agathosma ovata 'Glentana'	Soil	K25 AgS	***Kribbella sancticallisti* BC633(T)**	99.4
Agathosma ovata 'Glentana'	Soil	K40 AgS	***Kribbella sancticallisti* BC633(T)**	99.4
Tokai forest	Soil	P2A13	*Sporichthya brevicatena* IFO 16195(T)	95.4
Tokai forest	Soil	P2B10	*Actinoplanes tereljensis* MN07-A0371(T	97.4
Tokai forest	Soil	P2B20	*Sporichthya brevicatena* IFO 16195(T)	96.4
Tokai forest	Soil	S2A4	***Kribbella amoyensis* XMU 198(T)**	99.2
Tokai forest	Soil	S2A10	***Kribbella amoyensis* XMU 198(T)**	99.0

A breakdown of the sequence results among the plant samples revealed little difference in the diversity of strains between different plants and little difference between the leaf and root samples. Figure 3.1a-e illustrates the breakdown of sequences identified as *Kribbella* originating from either the leaf or the root material, as well as any sequences identified as non-*Kribbella* for each plant. Figure 3.1f represents an average of all the plants combined. It was initially expected that more *Kribbella* sequences would be present in the root samples, as the roots are in direct contact with the soil from which endophytes are acquired [Janso and Carter, 2010]. Grouping all the sequences by source (i.e. sample material), it can be seen in Figure 3.1f that 56% of all *Kribbella* sequences were from the leaf material, while 41% were from the root samples. The ratio of *Kribbella* sequences from leaf versus root varied with each plant sample, however the majority of the plant samples had a larger proportion of *Kribbella* sequences in the leaf samples as compared to the root samples (*L. conocarpodendron*: 56% leaf, 44% root; *A. ovata* 'Glentana': 69% leaf, 26% root; *P. aurea* subsp. *aurea*: 72% leaf, 14% root). The exceptions were *P. sericea*, which had a higher percentage in the root sample (81%) than the leaf sample (13%) (Figure 3.1a), and *O. frutescens*, which had an even representation of *Kribbella* sequences in the roots (48.5%) and leaves (48.5%) (Figure 3.1d).

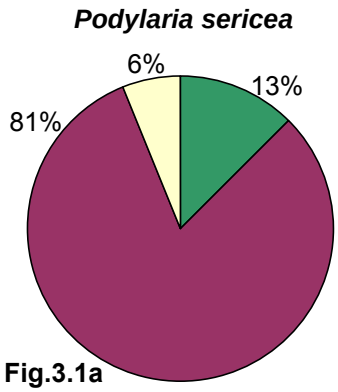


Fig.3.1a

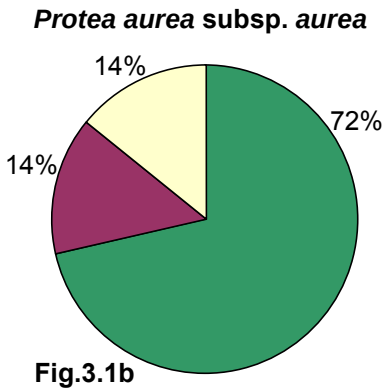


Fig.3.1b

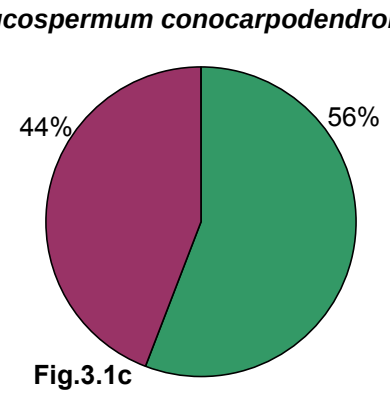


Fig.3.1c

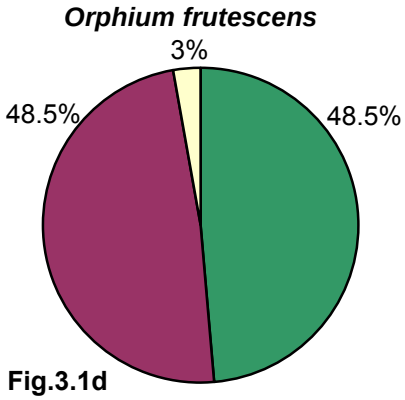


Fig.3.1d

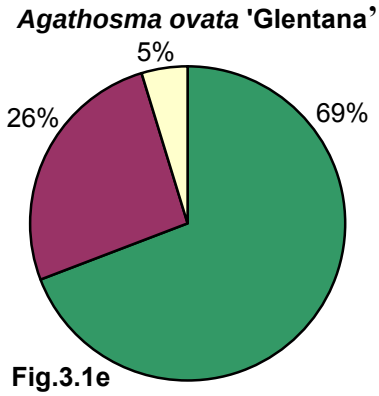


Fig.3.1e

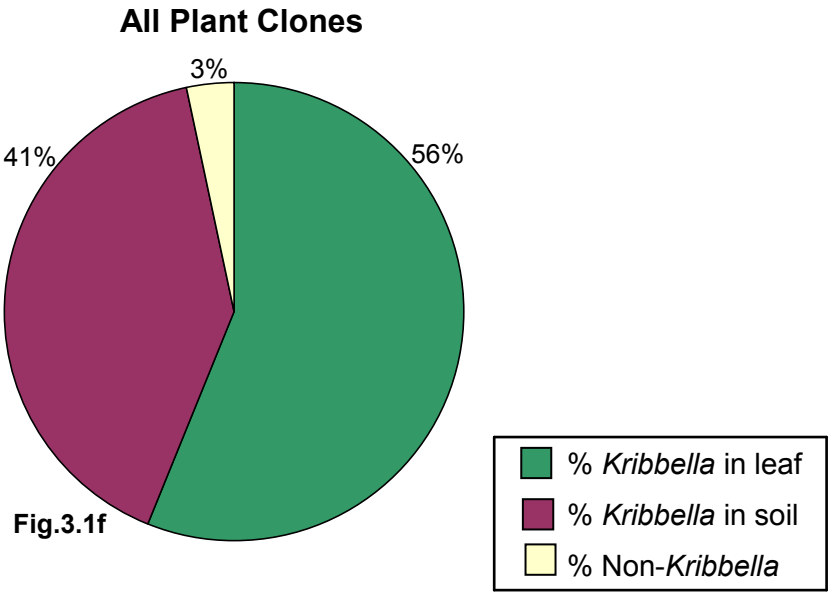


Fig.3.1f

Figure 3.1a-f. Clone sequences identified as being most similar to cultured type strains. The pie charts show the percentage of *Kribbella* sequences originating from leaf and root samples of each plant, as well as the percentage of non-*Kribbella* sequences isolated from each plant. a) *Podylaria sericea,* b) *Protea aurea* subsp. *aurea*, c) *Leucospermum conocarpodendron,* d) *Orphium frutescens,* e) *Agathosma ovata* 'Glentana', f) total of all plant clones.

As a further point of interest, the same metagenomic method was applied to the *Agathosma* soil sample using *gyrB* gene PCR primers GgyrB-F2 (GF2) and KgyrB-R1892 (K1892R) [Kirby *et al.,* 2010], as well as primers designed to amplify a portion of the *Kribbella rpoB* gene [Curtis and Meyers, 2012]. The *rpoB* gene primers (rpoBK1750F and rpoBK2292R) are described in Chapter 4. It was found that both sets of primers were able to amplify the appropriate genes from the soil sample. However, as can be seen in Table 3.2, the primers were not very specific for *Kribbella* species and all clone sequences had a low sequence similarity to their closest matches. Only one of the ten *gyrB* sequences had *Kribbella* as the top hit. Three *gyrB*-gene clones (SgB4, SgB6, SgB9) and one *rpoB*-gene clone (SrB52) did have *Kribbella* sequences listed in the BLAST results, but not as the top hit. Although, the *gyrB* and *rpoB* gene primers were able to amplify the correct gene from the soil sample, they were not specific for *Kribbella* sequences, as they were not designed to be specific to the genus, as the 16S rRNA gene primer Krb977F was. Nevertheless, all of the clone sequences belonged to actinobacteria (with the exception of those related to *Xanthomonas* and *Bradyrhizobium*).

Table 3.2: *gyrB* and *rpoB* gene clone sequence results from the GenBank database

Gene	Clone	GenBank top hit	% similarity
gyrB	SgB1	*Mycobacterium gilvum* Spyr1	86
gyrB	SgB2	*Mycobacterium indicus pranii* MTCC 9506	91
gyrB	SgB3	*Saccharothrix espanaensis* DSM 44229	86
gyrB	SgB4	*Streptomyces cattleya* DSM 46488	84
gyrB	SgB5	*Mycobacterium intracellulare* MOTT-64	88
gyrB	SgB6	*Frankia* sp. EAN1pec	82
gyrB	SgB7	***Kribbella aluminosa* HKI0480**	81
gyrB	SgB8	*Mycobacterium smegmatis* JS623	90
gyrB	SgB9	*Modestobacter marinus* BC501	91
gyrB	SgB10	*Xanthomonas albilineans* GPE PC73	77
rpoB	SrB10	*Cellulomonas fimi* ATCC 484	81
rpoB	SrB15	*Conexibacter woesei* DSM 14684	88
rpoB	SrB51	*Streptosporangium roseum* DSM 43021	88
rpoB	SrB52	*Thermobispora bispora* DSM 43833	93
rpoB	SrB57	*Bradyrhizobium japonicum* USDA 6	88

3.5 Discussion

The aim of this part of the study was to design a 16S rRNA gene PCR primer specific for the genus *Kribbella*. This primer was used to screen environmental samples for the presence of *Kribbella* strains. The Krb977F primer proved successful in amplifying *Kribbella* 16S rRNA gene sequences from two soil samples, as well as five leaf and five root samples. The results from these experiments indicate that *Kribbella* is present in a variety of plant samples and is not solely restricted to soil habitats.

The current 16S rRNA gene sequence similarity threshold for delineating a species is 98.7%. If the sequence similarity is below this threshold, DNA-DNA hybridization (DDH) is not necessary to establish that the two strains belong to distinct species [Stackebrandt and Ebers, 2006; Tindall *et al*., 2010]. However, a higher threshold of 99% 16S rRNA gene sequence similarity has recently been proposed as the threshold below which DDH is not required to distinguish between species [Meier-Kolthoff *et al.*, 2013]. However, it is important to note that not all bacterial taxonomists agree with these threshold proposals, preferring to use the widely accepted 16S rRNA gene sequence similarity threshold of 97% for distinguishing between genomic species without DDH analysis [Stackebrandt and Goebel, 1994]. These threshold values were based on comparisons of full length 16S rRNA gene sequences with DDH values and, although the sequences generated by the Krb977F/R5 primers are quite short (535 bp), the sequences highlighted in orange in Table 3.1 indicate potentially new *Kribbella* species.

Although, the low sequence similarities of some of the amplified environmental sequences indicate that the plant material could contain new *Kribbella* species, no *Kribbella* colonies were obtained from the coincident culture-dependant analysis of the plants. Interestingly, the Tokai forest soil sample that produced six strains of *K. solani* (Chapter 2) was not found to contain any *K. solani* sequences when screened by PCR with primers Krb977F/R5. The Tokai forest soil sample produced 25 clones, only five of which were identified as cultured type strains, two of which were identified as *K. amoyensis*. However, *K. solani* was one of the more prevalent strains found in the screening of the plant samples and a number of clones showed a 100% sequence similarity to *K. solani*. Since *K. solani* colonies were isolated from the Tokai forest soil,

the lack of detection of *K. solani* DNA in this soil sample could simply be the result of the low number of clones obtained from this sample.

It was also interesting to find that so many sequences originated from the leaf material, as it was initially expected that more *Kribbella* sequences would be present in the root samples, as the roots are in direct contact with the soil, from which endophytes are acquired [Janso and Carter 2010]. Although the overall number of sequences was higher in the leaf samples, the roots showed a slightly higher diversity of *Kribbella* species: nine species were detected in the roots versus only seven in the leaves. Previous metagenomic studies have found that relative to stems and leaves, roots harboured more diverse actinobacterial populations [Janso and Carter, 2010; Qin *et al.,* 2011]. The results of this study also agree with the results obtained by Bodenhausen *et al.*, [2013] that leaf and root endophytic compartments do not differ in richness and share many bacterial species.

This study also provides metagenomic sequence evidence that *Kribbella* can live endophytically in a variety of plants and plant parts. Most *Kribbella* strains have been isolated from soil or rock surfaces. However, the type strain of *Kribbella lupini* was isolated from roots and *K. endophytica* and '*K. podocarpi*' were isolated from leaves [Curtis *et al.,* manuscript under revision; Kaewkla and Franco, 2013; Trujillo *et al*., 2006]. The results from Figure 3.1f suggest that *Kribbella* strains have an equal affinity for roots and leaf tissue.

The data from this study suggest that *Kribbella* strains are ubiquitous in soil and in the leaves and roots of plants. The PCR-screening method developed can be utilized to screen soil and plant environmental samples for the presence of *Kribbella* strains before culture-dependent methods are used. The fact that this study established the presence of *Kribbella* endophytes in all the plants screened provides motivation to search for and isolate more endophytic *Kribbella* species. Plants may be an untapped environment for rarer genera of actinobacteria and this system of screening can be utilized and applied to other genera in the search for new secondary metabolites. Although *Kribbella* strains were readily detected in plant leaves and roots using the *Kribbella*-specific Krb977F primer, encouraging these endophytic strains to grow on agar isolation plates proved to

be significantly more difficult. This is an area for future investigation with different nutrient media and perhaps also media with reduced nutrient concentrations.

3.6 References

Bodenhausen, N., Horton, M.W. and Bergelson, J. (2013). Bacterial Communities Associated with the Leaves and the Roots of *Arabidopsis thaliana*. *PLoS ONE.* **8,** e56329. doi:10.1371/journal.pone.0056329

Cook, A.E. and Meyers, P. R. (2003) Rapid identification of filamentous actinomycetes to the genus level using genus-specific 16S-rRNA gene restriction fragment patterns. *Int. J. Syst. Evol. Microbiol*. **53**, 1907-1915.

Culligan, E.P., Sleator, R.D., Marchesi, J.R., and Hill, C. (2014). Metagenomics and novel gene discovery: Promise and potential for novel therapeutics. *Virulence*. **5**, 399-412.

Curtis, S.M., Norton, I., Everest, G.J. and Meyers, P.R. (2015). *Kribbella podocarpi* sp. nov., isolated from the leaves of a yellowwood tree (*Podocarpus latifolius*) and *Kribbella capetownensis* sp. nov. and *Kribbella speibonae* sp. nov., isolated from soil. *Antonie van Leeuwenhoek.* Manuscript under revision.

Gerhardt, P., Murray, R.G.E., Wood, W.A. and Krieg, N.R. Krieg**.** (1994). Methods for General and Molecular Bacteriology. ASM Press, Washington, D.C.

Graessler, J., Qin, Y., Zhong, H., Zhang, J., Licinio, J., Wong, M.-L., Xu, A., Chavakis, T., Bornstein, A.B., Erhart-Bornstein, M., Lamounier-Zepter, V., Lohmann, T., Wolf, T., and Bornstein S.R. (2013). Metagenomic sequencing of the human gut microbiome before and after bariatric surgery in obese patients with type 2 diabetes: correlation with inflammatory and metabolic parameters. *Pharmacogenomics*. **13**, 514-522.

Janso, J.E. and Carter, G.T. (2010). Biosynthetic Potential of Phylogenetically Unique Endophytic Actinomycetes from Tropical Plants. *Appl. Environ. Microbiol*. **76**, 4377-4386.

Kaewkla, O. and Franco, C.M.M. (2013). *Kribbella endophytica* sp. nov., an endophytic actinobacterium isolated from the surface-sterilized leaf of a native apricot tree. *Int. J. Syst. Evol. Microbiol.* **63**, 1249-1253.

Kim, O.S., Cho, Y.J., Lee, K., Yoon, S.H., Kim, M., Na, H., Park, S.C., Jeon, Y.S., Lee, J.H., Yi, H., Won, S., Chun, J. (2012). Introducing EZ-Taxon: a prokaryotic 16S rRNA Gene sequence database with phylotypes that represent uncultured species. *Int J Syst Evol Microbiol.* **62**, 716–721.

Kirby, B.M., Everest, G.J., and Meyers, P.R. (2010). Phylogenetic analysis of the genus *Kribbella* based on the *gyrB* gene: proposal of a *gyrB*-sequence threshold for species delineation in the genus *Kribbella*. *Antonie van Leeuwenhoek*. **97**, 131-142.

Kirby, B.M. and Meyers, P.R. (2010). *Micromonospora tulbaghiae* sp. nov., isolated from the leaves of wild garlic, *Tulbaghia violacea*. *Int. J. Syst. Evol. Microbiol.* **60**, 1328-1333.

Kurtböke, D.Í., (2012). Biodiscovery from rare actinomycetes: an eco-taxonomical perspective. *Appl. Microbiol. Biotechnol*. **93**, 1843-1852.

Lennox, E.S. (1955). Transduction of linked genetic characters of the host by bacteriophage P1. *Virology.* **1**,190-206.

Lepage, P., Leclerc, M.C., Joossens, M., Mondot, S., Blottière, H.M., Raes, J., Ehrlich, D. and Doré, J. (2013). A metagenomic insight into our gut's microbiome. *Gut*. **62**, 146-158.

Levy, R., and Borenstein, E. (2014). Metagenomics systems biology and metabolic modeling of the human microbiome. *Gut Microbes*. **5**, 1-6.

Meier-Kolthoff, J. P., Göker, M., Spröer, C. and Klenk, H. P. (2013). When should a DDH experiment be mandatory in microbial taxonomy? *Arch. Microbiol.* **195**, 413-418.

Messing, J. (1983). New M13 vectors for cloning. *Methods in Enzymology*. **101**, 20–78.

Qin, S., Li, J., Chen, H.-H., Zhao, G.-Z., Zhu, W.-Y., Jiang, C.-L., Xu, L.-H. and Li, W.-J. (2009). Isolation, Diversity, and Antimicrobial Activity of Rare Actinobacteria from Medicinal Plants of Tropical Rain Forests in Xishuangbanna, China. *Appl. Environ Microbiol.* **75**, 6176-6186.

Qin, S., Xing, K., Jiang, J.-H., Xu, L.-H. and Li, W.-J. (2011) Biodiversity, bioactive natural products and biotechnological potential of plant-associated endophytic actinobacteria. *Appl. Microbiol. Biot*. **89**, 457-473.

Raes, J. (2014). The gut microbiome- a new target for understanding, diagnosing and treating disease. *Arch. Public Health.* **72,** Suppl 1-K3.

Riesenfeld, C.S., Schloss, P.D., and Handelsman, J. (2004). Metagenomics: Genomic Analysis of Microbial Communities. *Annu. Rev. Genet*. **38**, 525-552.

Rondon, M.R., August, P.R., Betterman, A.D., Brady, S.F., Grossman, T.H., Liles, M.R., Loiacono, K.A., Lynch, B.A., MacNeil, I.A., Minor, C., Tiong, C.L., Gilman, M., Osburne, M.S., Clardy, J., Handelsman, J. and Goodman, R.M. (2000). Cloning the Soil Metagenome: a Strategy for Accessing the Genetic and Functional Diversity of Uncultured Microorganism. *Appl. Environ. Microbiol*. **66**, 2541-2547.

Schloss, P.D. and Handelsman, J. (2003). Biotechnological prospects from metagenomics. *Curr. Opin. Biotech*. **14**, 303-310.

Schmeisser, C., Steel, H. and Streit, W.R. (2007). Metagenomics, biotechnology with non-culturable microbes. *Appl. Microbiol. Biot*. **75**, 955-962.

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. and Ebers, J. (2006). Taxonomic parameters revisited: tarnished gold standards. *Microbiol. Today*. **33**, 152-1155.

Strobel, G., Daisy, B., Castillo, U., and Harper, J. (2004). Natural products from endophytic microorganisms. *J. Nat. Prod*. **67**, 257-268.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol*. **28**, 273-2739.

Tindall, B.J., Rosselló-Móra, R., Busse, H.-J., Ludwig, W. and Kämpfer, P. (2010). Notes on the characterization of prokaryote strains for taxonomic purposes. *Int. J. Syst. Evol. Microbiol*. **60**, 249-266.

Trujillo, M.E., Kroppenstedt, R.M., Schumann, P., and Martinez-Molina, E. (2006). *Kribbella lupini* sp. nov., isolated from the roots of *Lupinus angustifolius*. *Int. J. Syst. Evol. Microbiol*. **56**, 407-411.

Turnbaugh, P.J., Ridaura, V.K., Faith, J.J., Rey, F.E., Knight, R. and Gordon J.I. (2009). The Effect of Diet on the Human Gut Microbiome: A Metagenomci Analysis in Humanized Gnotobiotic Mice. *Sci. Transl. Med*. **1**, 6ra14.

3.7 Appendices

Appendix 3A: List of strains and accession numbers used in the alignment to design the *Kribbella*-specific 16S rRNA gene primer.

Strains used in *Nocardioides* consensus sequence	
Strain	**Accession Number**
Nocardioides aestuarii JC2056	AY423719
Nocardioides agariphilus MSL 28	EF466113
Nocardioides albus KCTC 9186	AF004988
Nocardioides alkalitolerans KSL-1	AY633969
Nocardioides aquaticus DSM 11439	X94145
Nocardioides aquiterrae GW-9	AF529063
Nocardioides aromaticivorans H-1	AB087721
Nocardioides basaltis J112	EU143365
Nocardioides bigeumensis MSL 19	EF466114
Nocardioides caeni MN8	FJ423551
Nocardioides carcicola YC6903	FJ750845
Nocardioides daedukensis MDN22	FJ842646
Nocardioides daphnia D287T	AM398438
Nocardioides dilutus MSL 11	EF466121
Nocardioides dokdonensis FR1436	EF633986
Nocardioides exalbidus RC825	AB273624
Nocardioides fastidiosum JCM 8088	AF005022
Nocardioides fonticola NAA-13	EF626689
Nocardioides furvisabuli SBS-26	DQ411542
Nocardioides ganghwensis JC2055	AY423718
Nocardioides ginsengisoli DSM 17921	AB245396.
Nocardioides halotolerans MSL-23	EF466122
Nocardioides hankookensis DS-30	EF555584
Nocardioides humi DCY24	EF623863
Nocardioides hwasunensis HFW-21	AM295258
Nocardioides insulae DS-51	DQ786794
Nocardioides islandensis MSL 26	EF466123
Nocardioides jensenii ATCC 49810	AF005006
Nocardioides kongjuensis A2-4	DQ218275
Nocardioides koreensis MSL 09	EF466115
Nocardioides kribbensis KSL-2	AY835924
Nocardioides lentus KSL-17	DQ121389
Nocardioides luteus KCTC 9575	AF005007
Nocardioides marinisabuli SBS-12T	AM422448
Nocardioides marinus CL-DD14	DQ401093
Nocardioides mesophlius MSL 22	EF466117
Nocardioides nitrophenolicus NSP 41	AF005024
Nocardioides panacihumi DSM 18660	AB271053

Strains used in alignment for primer design not of the *Nocardioides* genus	
Strain	**Accession Number**
Actinopolymorpha alba YIM48868	EU706350
Actinopolymorpha cephalotaxi DSM 4511	EU438909
Actinopolymorpha rutila YIM 45725	EF601829.
Actinopolymorpha singaporensis JCM 10	AF237815
Aeromicrobium alkaliterrae DSM 16824	AY822044
Aeromicrobium erythreum NRRL B-3381	AF005021
Aeromicrobium fastidiosum DSM 10552	AF005022
Aeromicrobium flavum TYLN1	EF133690
Aeromicrobium ginsengisoli JCM 14732	AB245394
Aeromicrobium halocynthiae KME 001	FJ042789
Aeromicrobium marinum DSM 15272	AY166703
Aeromicrobium panaciterrae DSM 17939	AB245387
Aeromicrobium ponti HSW-1T	AM778683
Aeromicrobium tamlense SSW1-57	DQ411541
Jiangella alba YIM 6150	FJ157186
Jiangella alkaliphila D8-87T	AM422451
Jiangella gansuensis YIM 002	AY631071
Jiangella muralis DSM 44835	FN645214
Marmoricola aequoreus SST-45T	AM295338
Marmoricola aurantiacus DSM 12652	Y18629
Marmoricola bigeumensis MSL 05	EF466120
Marmoricola scoriae Sco-D01	FN386750
Aestuariimicrobium kwangyangense R27	DQ830982
Brooklawnia cerclae BL-34	DQ196625
Friedmanniella antarctica DSM 11053	AF110052
Friedmanniella capsulata ACM 5120	AF084529
Friedmanniella lacustris EL-17a	AJ132943
Friedmanniella lucida DSM 21742	AB445454
Friedmanniella luteola FA1	AB445453
Friedmanniella okinawensis FB1	AB445455
Friedmanniella sagamiharensis FB2	AB445456
Friedmanniella spumicola ACM 5121	AF062535
Granulicoccus phenolivorans PG-02	AY566575
Luteococcus japonicus ATCC 51526	D85487
Luteococcus peritonei ATCC BAA-60	AJ132334
Luteococcus sanguinis CCUG 33897	AJ416758
Microlunatus aurantiacus YIM 45721	EF601828
Microlunatus panaciterrae DSM 18662	AB271051

Appendix 3A continued

Strains used in *Nocardioides* consensus sequence	
Strain	**Accession Number**
Nocardioides panacisoli DSM 21348	FJ666101
Nocardioides plantarum NCIMB 12834	AF005008.
Nocardioides pyridinolyticus OS4	U61298
Nocardioides salarius CL-Z59	DQ401092
Nocardioides sediminis MSL-01	EF466110
Nocardioides simplex KCTC 9106	AF005009
Nocardioides terrae VA15	FJ423762
Nocardioides terrigena DS-17	EF363712
Nocardioides tritolerans MSL-14v	EF466107

Strains used in alignment for primer design not of the *Nocardioides* genus	
Strain	**Accession Number**
Microlunatus parietis DSM 22083	FN556016
Microlunatus phosphovorus NM-1	Z78207
Microlunatus soli CC-012602	FJ807672
Micropruina glycogenica Lg2	AB012607
Propionibacterium acidifaciens C3M_31	EU979537
Propionibacterium acidipropionici NCFB 570	AJ704570
Propionibacterium acidipropionici NCFB 563	AJ704569
Propionibacterium australiense LCDC-9	AF225962
Propionibacterium avidum DSM 4901	AJ003055
Propionibacterium cyclohexanicum TA-1	NR_036827
Propionibacterium freudenreichii ATCC	AY533300
Propionibacterium freudenreichii subsp. shermanii	HM626365
Propionibacterium granulosum DSM 2070	KF906605
Propionibacterium jensenii DSM 20535	AJ704571
Propionibacterium lymphophilum DSM 49	AJ003056
Propionibacterium microaerophilus M5	AF234623.
Propionibacterium propionicus DSM 433	AF285117
Propionibacterium thoenii NCFB568	AJ704572
Propionicicella superfundia BL-10T	DQ176646
Propionicimonas paludicola DSM 15597	FR733712
Propioniferax innocua L60	AF227165
Tessaracoccus bendigoniensis ACM 5119	AF038504.
Tessaracoccus flavescens SST-39T	AM393882
Tessaracoccus lubricantis KSS-17SeT	FM178840

Appendix 3B: List of the non-*Kribbella* strains from which the primer pair F1/Krb624R amplified a PCR product. For all strains tested, Krb977F/R5 did not produce a product.

Genera and species tested
Herbidospora yilanensis NBRC 106371^T
Nonomuraea africana NRRL B-16114^T
Nonomuraea angiospora NRRL B-3905^T
Nonomuraea antimicrobica DSM 45220^T
Nonomuraea bangladeshensis DSM 45128^T
Nonomuraea fastidiosa NBRC 14680^T
Nonomuraea ferruginea NRRL B-16096^T
Nonomuraea helvata NRRL B-16123^T
Nonomuraea kuesteri NRRL B-24325^T
Nonomuraea longicatena NBRC 16462^T
Nonomuraea maheshkhaliensis DSM 45163^T
Nonomuraea polychroma NRRL B-16243^T
Nonomuraea pusilla NRRL B-16126^T
Nonomuraea recticatena NBRC 14525^T
Nonomuraea roseola NBRC 14685^T
Nonomuraea rubra NRRL B-16083^T
Nonmuraea salmonea NBRC 14687^T
Nonomuraea spiralis NBRC 14097^T
Nonomuraea turkmeniaca NRRL B-16246^T
Planomonospora sphaerica DSM 44632^T
Sphaerisporangium album DSM 45172^T
Sphaerisporangium flaviroseum DSM 45170^T
Sphaerisporangium viridialbum NRRL B-2636^T
Streptosporangium album NRRL B-2635^T
Streptosporangium amethystogenes NRRL B-2639^T
Streptosporangium carneum NBRC 15562^T
Streptosporangium fragile NRRL B-16437^T
Streptosporangium nondiastaticum DSM 43848^T
Streptosporangium pseudovulgare NBRC 13991^T
Streptosporangium violaceochromogenes NRRL B-16784^T
Streptosporangium vulgare NRRL B-2633^T

Appendix 3C: Ez-Taxon-e and GenBank BLAST results for all metagenomic clones

Plant Name	Plant material	Clone Name	Ez-Taxon-e Top Hit	% Similarity	GenBank Top hit	% Similarity
Podylaria sericea	Leaf	PodsL15	Uncultured Acidobacteria GQ397046_s AK4DE1_07H	95.3	Uncultured actinobacterium	99
Podylaria sericea	Leaf	PodsL16	*Ilumatobacter fluminis YM22-133(T)*	96.6	Uncultured bacterium	98
Podylaria sericea	Leaf	PodsL25	***Kribbella hippodromi S1.4(T)***	98.6	***Kribbella solani* DSA1(T)**	98
Podylaria sericea	Leaf	PodsL34	***Kribbella sancticallisti BC633(T)***	99.2	***Kribbella aluminosa* HKI 0480**	99
Podylaria sericea	Leaf	PodsL43	Uncultured *Chthoniobacter* AJ937874_s A825.05.04, 9m	51.6	Cloning vector pMono_T-vector	100
Podylaria sericea	Leaf	PodsL66	Uncultured *Ilumatobacter* AF418962_s HTG5	51.5	Uncultured bacterium	98
Podylaria sericea	Root	PodsR3	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	96.2	Uncultured bacterium	97
Podylaria sericea	Root	PodsR5	***Kribbella aluminosa* HKI 0478(T)**	99.8	***Kribbella aluminosa* HKI 0480**	99
Podylaria sericea	Root	PodsR6	***Kribbella hippodromi* S1.4(T)**	99.0	***Kribbella solani* DSA1(T)**	99
Podylaria sericea	Root	PodsR8	***Kribbella hippodromi* S1.4(T)**	99.0	***Kribbella solani* DSA1(T)**	99
Podylaria sericea	Root	PodsR43	***Kribbella solani* DSA1(T)**	98.3	***Kribbella solani* DSA1(T)**	98
Podylaria sericea	Root	PodsR45	***Kribbella karoonensis* DSM 17344(T)**	97.3	***Kribbella solani* DSA1(T)**	95
Podylaria sericea	Root	PodsR46	***Kribbella hippodromi* S1.4(T)**	99.2	***Kribbella solani* DSA1 (T)**	99
Podylaria sericea	Root	PodsR49	***Kribbella solani* DSA1(T)**	98.1	***Kribbella solani* DSA1(T)**	98
Podylaria sericea	Root	PodsR50	***Kribbella karoonensis* DSM 17344(T)**	98.5	***Kribbella solani* DSA1(T)**	97
Podylaria sericea	Root	PodsR58	***Kribbella hippodromi* S1.4(T)**	99.6	***Kribbella solani* DSA1 (T)**	99
Podylaria sericea	Root	PodsR59	***Kribbella karoonensis* DSM 17344(T)**	97.3	***Kribbella aluminosa* HKI 0480**	93
Podylaria sericea	Root	PodsR60	***Kribbella hippodromi* S1.4(T)**	99.4	***Kribbella solani* DSA1(T)**	99
Podylaria sericea	Root	PodsR62	***Kribbella karoonensis* DSM 17344(T)**	95.3	***Kribbella hippodromi* S1.4(T)**	95
Podylaria sericea	Root	PodsR73	***Kribbella hippodromi* S1.4(T)**	99.2	***Kribbella solani* DSA1(T)**	99
Protea aurea subsp. *aurea*	Leaf	ProtaL3	Uncultured *Gemmatimonadetes* EU134894_s FFCH6442	90.2	Uncultured bacterium	93
Protea aurea subsp. *aurea*	Leaf	ProtaL23	***Kribbella sancticallisti* BC633(T)**	99.0	***Kribbella solani* DSA1(T)**	98
Protea aurea subsp. *aurea*	Leaf	ProtaL32	***Kribbella sancticallisti* BC633(T)**	99.2	***Kribbella solani* DSA1(T)**	99
Protea aurea subsp. *aurea*	Leaf	ProtaL42	***Kribbella hippodromi* S1.4(T)**	99.6	***Kribbella solani* DSA1(T)**	99
Protea aurea subsp. *aurea*	Leaf	ProtaL49	***Kribbella sancticallisti* BC633(T)**	99.2	***Kribbella solani* DSA1(T)**	99
Protea aurea subsp. *aurea*	Leaf	ProtaL57	Uncultured *Acidimicrobiales* FJ478903_s p11i04ok	94.7	Uncultured bacterium	97
Protea aurea subsp. *aurea*	Leaf	ProtaL66	Uncultured *lamiaceae* FM209118_s delph1F3	96.9	Uncultured bacterium	98
Protea aurea subsp. *aurea*	Leaf	ProtaL71	***Kribbella karoonensis* DSM 17344(T)**	99.5	***Kribbella aluminosa* HKI 0480**	99

Appendix 3C continued

Plant Name	Plant material	Clone Name	Ez-Taxon-e Top Hit	% Similarity	GenBank Top hit	% Similarity
Protea aurea subsp. *aurea*	Root	ProtaR3	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	97.9	Uncultured bacterium	98
Protea aurea subsp. *aurea*	Root	ProtaR5	Uncultured *Iamiaceae* FM209069_s delph1B1	96.9	Uncultured bacterium	97
Protea aurea subsp. *aurea*	Root	ProtaR6	Uncultured *Ilumatobacter* EU799964_s 1C227642	96.7	Uncultured bacterium	97
Protea aurea subsp. *aurea*	Root	ProtaR7	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	97.7	Uncultured bacterium	98
Protea aurea subsp. *aurea*	Root	ProtaR8	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	97.8	Uncultured bacterium	98
Protea aurea subsp. *aurea*	Root	ProtaR9	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	95.9	Uncultured bacterium	98
Protea aurea subsp. *aurea*	Root	ProtaR10	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	97.9	Uncultured bacterium	98
Protea aurea subsp. *aurea*	Root	ProtaR11	Uncultured *Iamiaceae* FM209069_s delph1B1	97.6	Uncultured bacterium	97
Protea aurea subsp. *aurea*	Root	ProtaR13	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	97.9	Uncultured bacterium	98
Protea aurea subsp. *aurea*	Root	ProtaR14	Uncultured *Ilumatobacter* FJ358904_s OX26	96.5	Uncultured actinobacterium	96
Protea aurea subsp. *aurea*	Root	ProtaR18	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	97.8	Uncultured bacterium	98
Protea aurea subsp. *aurea*	Root	ProtaR19	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	97.2	Uncultured bacterium	98
Protea aurea subsp. *aurea*	Root	ProtaR20	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	97.7	Uncultured bacterium	98
Protea aurea subsp. *aurea*	Root	ProtaR22	Uncultured *Ilumatobacter* AY532590_s	95.5	Uncultured *Acidimicrobiales* bacterium	95
Protea aurea subsp. *aurea*	Root	ProtaR23	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	97.7	Uncultured bacterium	98
Protea aurea subsp. *aurea*	Root	ProtaR30	Uncultured *Ilumatobacter* FJ358904_s OX26	96.6	Uncultured bacterium	96
Protea aurea subsp. *aurea*	Root	ProtaR31	Uncultured *Iamiaceae* EU799964_s 1C227642	96.7	Uncultured bacterium	97
Protea aurea subsp. *aurea*	Root	ProtaR33	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	97.9	Uncultured bacterium	98
Protea aurea subsp. *aurea*	Root	ProtaR37	Uncultured *Ilumatobacter* FN396713_s s5_8_III_16	94.5	Uncultured bacterium	97
Protea aurea subsp. *aurea*	Root	ProtaR41	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	97.7	Uncultured bacterium	98
Protea aurea subsp. *aurea*	Root	ProtaR42	***Kribbella hippodromi* S1.4(T)**	99.6	***Kribbella solani* DSA1(T)**	99
Protea aurea subsp. *aurea*	Root	ProtaR46	Uncultured Iamiaceae FM209069_s delph1B1	97.2	Uncultured bacterium	97
Protea aurea subsp. *aurea*	Root	ProtaR48	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	97.3	Uncultured bacterium	98
Protea aurea subsp. *aurea*	Root	ProtaR50	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	97.7	Uncultured bacterium	98
Protea aurea subsp. *aurea*	Root	ProtaR51	Uncultured *Ilumatobacter* FJ358904_s OX26	96.3	Uncultured actinobacterium	96
Protea aurea subsp. *aurea*	Root	ProtaR54	Uncultured *Acidimicrobiales* EU374087_s HCM3MC78_8D_FL	71.7	Uncultured bacterium	94
Protea aurea subsp. *aurea*	Root	ProtaR56	Uncultured *Iamiaceae* FM209138_s delph2G1	96.9	Uncultured bacterium	98
Protea aurea subsp. *aurea*	Root	ProtaR61	Uncultured *Acidimicrobiales* EU491412_s P0X4b3E02	50.2	Uncultured bacterium	98
Protea aurea subsp. *aurea*	Root	ProtaR62	Uncultured *Iamiaceae* FJ152827_s TX5A_119	95.2	Uncultured bacterium	95
Protea aurea subsp. *aurea*	Root	ProtaR65	*Ilumatobacter fluminis* YM22-133(T)	96.8	Uncultured bacterium	98

Appendix 3C continued

Plant Name	Plant material	Clone Name	Ez-Taxon-e Top Hit	% Similarity	GenBank Top hit	% Similarity
Protea aurea subsp. *aurea*	Root	ProtaR70	Uncultured *Ilumatobacter* AY532590_s 1013-28-CG55	97.6	Uncultured bacterium	98%
Leucospermum conocarpodendron	Leaf	LcL1	***Kribbella karoonensis* Q41(T)**	92.4	***Kribbella sandramycini* ATCC 39419(T)**	91
Leucospermum conocarpodendron	Leaf	LcL2	***Kribbella solani* DSA1(T)**	99.6	***Kribbella solani* DSA1(T)**	99
Leucospermum conocarpodendron	Leaf	LcL3	Uncultured *Gemmatimonadales* EF516935_s FCPN548	96.9	Uncultured bacterium	97
Leucospermum conocarpodendron	Leaf	LcL4	***Kribbella aluminosa* HKI 0478(T)**	99.8	***Kribbella aluminosa* HKI 0480**	99
Leucospermum conocarpodendron	Leaf	LcL5	***Kribbella aluminosa* HKI 0478(T)**	99.8	***Kribbella aluminosa* HKI 0480**	99
Leucospermum conocarpodendron	Leaf	LcL6	***Kribbella aluminosa* HKI 0478(T)**	99.4	***Kribbella aluminosa* XMU153**	99
Leucospermum conocarpodendron	Leaf	LcL7	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	97.7	Uncultured bacterium	98
Leucospermum conocarpodendron	Leaf	LcL8	Uncultured *Gemmatimonadales* EF516935_s FCPN548	96.7	Uncultured bacterium	97
Leucospermum conocarpodendron	Leaf	LcL9	Uncultured *Gemmatimonadales* EF516935_s FCPN548	96.9	Uncultured bacterium	97
Leucospermum conocarpodendron	Leaf	LcL10	Uncultured *Jatrophihabitans* AB245400_s Gsoil 1621	98.1	*Actinomycetales* bacterium Gsoil 1621	98
Leucospermum conocarpodendron	Leaf	LcL11	***Kribbella aluminosa* HKI 0478(T)**	99.8	***Kribbella aluminosa* HKI 0480**	99
Leucospermum conocarpodendron	Leaf	LcL12	***Kribbella sancticallisti* BC633(T)**	99.2	***Kribbella solani* DSA1(T)**	99
Leucospermum conocarpodendron	Leaf	LcL13	***Kribbella sancticallisti* BC633(T)**	99.2	***Kribbella solani* DSA1(T)**	99
Leucospermum conocarpodendron	Leaf	LcL14	Uncultured *Gemmatimonadales* EF516935_s FCPN548	96.7	Uncultured bacterium	97
Leucospermum conocarpodendron	Leaf	LcL15	***Kribbella solani* DSA1(T)**	99.6	***Kribbella solani* DSA1(T)**	99
Leucospermum conocarpodendron	Leaf	LcL16	***Kribbella karoonensis* DSM 17344(T)**	99.8	***Kribbella aluminosa* HKI 0480**	99
Leucospermum conocarpodendron	Leaf	LcL17	***Kribbella swartbergensis* HMC25(T)**	98.9	***Kribbella solani* DSA1(T)**	99
Leucospermum conocarpodendron	Leaf	LcL18	Uncultured *Gemmatimonadales* AJ532728_s JG34-KF-418	96.8	Uncultured bacterium	96
Leucospermum conocarpodendron	Leaf	LcL20	***Kribbella karoonensis* DSM 17344(T)**	93.0	Uncultured bacterium	91
Leucospermum conocarpodendron	Leaf	LcL21	***Kribbella sancticallisti* BC633(T)**	99.4	***Kribbella solani* DSA1(T)**	99
Leucospermum conocarpodendron	Leaf	LcL22	Uncultured *Iamiaceae* FM209118_s delph1F3	96.9	Uncultured bacterium	97
Leucospermum conocarpodendron	Leaf	LcL23	***Kribbella aluminosa* HKI 0478(T)**	99.6	***Kribbella aluminosa* HKI 0480**	99
Leucospermum conocarpodendron	Leaf	LcL24	Uncultured *Acidimicrobiales* GQ397046_s AK4DE1_07H	97.7	Uncultured actinobacterium	99
Leucospermum conocarpodendron	Leaf	LcL25	***Kribbella aluminosa* HKI 0478(T)**	99.8	***Kribbella aluminosa* HKI 0480**	99
Leucospermum conocarpodendron	Leaf	LcL27	Uncultured *Acidimicrobiales* GQ397046_s AK4DE1_07H	97.7	Uncultured actinobacterium	99
Leucospermum conocarpodendron	Leaf	LcL29	Uncultured *Acidimicrobiales* GQ397046_s AK4DE1_07H	95.2	Uncultured actinobacterium	99
Leucospermum conocarpodendron	Leaf	LcL30	***Kribbella sancticallisti* BC633(T)**	99.0	***Kribbella solani* DSA1(T)**	99
Leucospermum conocarpodendron	Leaf	LcL31	***Kribbella karoonensis* DSM 17344(T)**	99.5	***Kribbella aluminosa* HKI 0480**	96
Leucospermum conocarpodendron	Leaf	LcL33	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	100.0	***Kribbella solani* DSA1(T)**	99

Appendix 3C continued

Plant Name	Plant material	Clone Name	Ez-Taxon-e Top Hit	% Similarity	GenBank Top hit	% Similarity
Leucospermum conocarpodendron	Leaf	LcL34	Uncultured *Iamiaceae* FM209069_s delph1B1	97.4	Uncultured bacterium	97
Leucospermum conocarpodendron	Leaf	LcL35	Uncultured *Acidimicrobiales* DQ129383_s AKIW874	49.7	Uncultured bacterium	76
Leucospermum conocarpodendron	Leaf	LcL38	***Kribbella karoonensis* DSM 17344(T)**	99.8	***Kribbella aluminosa* HKI 0480**	99
Leucospermum conocarpodendron	Leaf	LcL39	***Kribbella karoonensis* DSM 17344(T)**	99.5	***Kribbella aluminosa* HKI 0480**	99
Leucospermum conocarpodendron	Leaf	LcL41	***Kribbella shirazensis* UTMC 693(T)**	81.4	***Kribbella aluminosa* HKI 0480**	80
Leucospermum conocarpodendron	Leaf	LcL42	***Kribbella karoonensis* DSM 17344(T)**	91.1	***Kribbella hippodromi* S1.4(T)**	88
Leucospermum conocarpodendron	Leaf	LcL44	***Kribbella aluminosa* HKI 0478(T)**	99.8	***Kribbella aluminosa* HKI 0480**	99
Leucospermum conocarpodendron	Leaf	LcL45	***Kribbella solani* DSA1(T)**	99.6	***Kribbella solani* DSA1(T)**	99
Leucospermum conocarpodendron	Leaf	LcL46	Uncultured *Iamiaceae* FM209118_s delph1F3	96.9	Uncultured bacterium	98
Leucospermum conocarpodendron	Leaf	LcL47	***Kribbella sancticallisti* BC633(T)**	99.4	***Kribbella solani* DSA1(T)**	99
Leucospermum conocarpodendron	Leaf	LcL48	***Kribbella sancticallisti* BC633(T)**	99.2	***Kribbella solani* DSA1(T)**	99
Leucospermum conocarpodendron	Leaf	LcL49	***Kribbella aluminosa* HKI 0478(T)**	99.8	***Kribbella aluminosa* HKI 0480**	99
Leucospermum conocarpodendron	Leaf	LcL50	***Kribbella solani* DSA1(T)**	99.6	***Kribbella solani* DSA1(T)**	99
Leucospermum conocarpodendron	Leaf	LcL53	***Kribbella aluminosa* HKI 0478(T)**	99.1	***Kribbella aluminosa* XMU153**	99
Leucospermum conocarpodendron	Leaf	LcL54	Uncultured *Acidimicrobiales* GQ397046_s AK4DE1_07H	97.7	Uncultured actinobacterium	99
Leucospermum conocarpodendron	Leaf	LcL55	Uncultured *Gemmatimonadales* EF516935_s FCPN548	96.9	Uncultured bacterium	97
Leucospermum conocarpodendron	Leaf	LcL56	***Kribbella aluminosa* HKI 0478(T)**	99.6	***Kribbella aluminosa* HKI 0480**	99
Leucospermum conocarpodendron	Leaf	LcL57	***Kribbella sancticallisti* BC633(T)**	99.4	***Kribbella solani* DSA1(T)**	99
Leucospermum conocarpodendron	Leaf	LcL58	Uncultured *Ilumatobacter* FJ358904_s OX26	95.2	Uncultured bacterium	95
Leucospermum conocarpodendron	Leaf	LcL59	Uncultured *Gemmatimonadales* EF516935_s FCPN548	96.9	Uncultured bacterium	97
Leucospermum conocarpodendron	Leaf	LcL60	***Kribbella aluminosa* HKI 0478(T)**	100.0	***Kribbella aluminosa* HKI 0480**	100
Leucospermum conocarpodendron	Leaf	LcL71	***Kribbella sancticallisti* BC633(T)**	99.2	***Kribbella solani* DSA1(T)**	99
Leucospermum conocarpodendron	Root	LcR1	***Kribbella aluminosa* HKI 0478(T)**	98.7	***Kribbella aluminosa* HKI 0480**	99
Leucospermum conocarpodendron	Root	LcR2	Uncultured *Gemmatimonadales* EF516935_s FCPN548	96.7	Uncultured bacterium clone	97
Leucospermum conocarpodendron	Root	LcR3	***Kribbella solani* DSA1(T)**	99.8	***Kribbella solani* DSA1(T)**	99
Leucospermum conocarpodendron	Root	LcR7	Uncultured *Gemmatimonadales* AJ532728_s JG34-KF-418	93.9	Uncultured bacterium	92
Leucospermum conocarpodendron	Root	LcR8	***Kribbella hippodromi* S1.4(T)**	100.0	***Kribbella solani* DSA1(T)**	99
Leucospermum conocarpodendron	Root	LcR11	***Kribbella aluminosa* HKI 0478(T)**	99.6	***Kribbella aluminosa* HKI 0480**	99
Leucospermum conocarpodendron	Root	LcR12	***Kribbella aluminosa* HKI 0478(T)**	100.0	***Kribbella aluminosa* HKI 0480**	100
Leucospermum conocarpodendron	Root	LcR14	***Kribbella aluminosa* HKI 0478(T)**	99.8	***Kribbella aluminosa* HKI 0480**	99

Appendix 3C continued

Plant Name	Plant material	Clone Name	Ez-Taxon-e Top Hit	% Similarity	GenBank Top hit	% Similarity
Leucospermum conocarpodendron	Root	LcR15	***Kribbella aluminosa* HKI 0478(T)**	99.6	***Kribbella aluminosa* HKI 0480**	99
Leucospermum conocarpodendron	Root	LcR18	Uncultured *Acidimicrobiales* GQ397046_s AK4DE1_07H	97.7	Uncultured actinobacterium	98
Leucospermum conocarpodendron	Root	LcR19	***Kribbella aluminosa* HKI 0478(T)**	99.8	***Kribbella aluminosa* HKI 0480**	99
Leucospermum conocarpodendron	Root	LcR20	***Kribbella aluminosa* HKI 0478(T)**	99.6	***Kribbella aluminosa* HKI 0480**	99
Leucospermum conocarpodendron	Root	LcR22	***Kribbella aluminosa* HKI 0478(T)**	99.6	***Kribbella aluminosa* HKI 0480**	99
Leucospermum conocarpodendron	Root	LcR23	***Kribbella aluminosa* HKI 0478(T)**	99.8	***Kribbella aluminosa* HKI 0480**	99
Leucospermum conocarpodendron	Root	LcR24	***Kribbella solani* DSA1(T)**	98.3	***Kribbella solani* DSA1(T)**	98
Leucospermum conocarpodendron	Root	LcR25	***Kribbella hippodromi* S1.4(T)**	100.0	***Kribbella solani* DSA1(T)**	99
Leucospermum conocarpodendron	Root	LcR26	Uncultured *Jatrophihabitans* AB245400_s Gsoil 1621	97.9	*Actinomycetales* bacterium	98
Leucospermum conocarpodendron	Root	LcR27	Uncultured *Iamiaceae* GQ487905_s T1-39	48.8	Uncultured bacterium	97
Leucospermum conocarpodendron	Root	LcR28	***Kribbella aluminosa* HKI 0478(T)**	100.0	***Kribbella aluminosa* HKI 0480**	100
Leucospermum conocarpodendron	Root	LcR29	Uncultured *Acidimicrobiales* GQ397046_s AK4DE1_07H	97.7	Uncultured actinobacterium	99
Leucospermum conocarpodendron	Root	LcR34	***Kribbella karoonensis* DSM 17344(T)**	99.5	***Kribbella aluminosa* HKI 0480**	99
Leucospermum conocarpodendron	Root	LcR35	***Kribbella sancticallisti* BC633(T)**	99.2	***Kribbella solani* DSA1(T)**	99
Leucospermum conocarpodendron	Root	LcR37	Uncultured *Gemmatimonadales* EF516935_s FCPN548	96.7	Uncultured bacterium	97
Leucospermum conocarpodendron	Root	LcR40	***Kribbella aluminosa* HKI 0478(T) / *Kribbella karoonensis* DSM 17344(T)**	100.0	***Kribbella aluminosa* HKI 0480**	100
Leucospermum conocarpodendron	Root	LcR41	Uncultured *Gaiella* JF429126_s C73	94.1	Uncultured actinobacterium	94
Leucospermum conocarpodendron	Root	LcR42	***Kribbella sancticallisti* BC633(T)**	99.0	***Kribbella solani* DSA1(T)**	99
Leucospermum conocarpodendron	Root	LcR45	***Kribbella aluminosa* HKI 0478(T)**	100.0	***Kribbella aluminosa* HKI 0480**	100
Leucospermum conocarpodendron	Root	LcR49	***Kribbella swartbergensis* DSM 17345(T)**	98.8	***Kribbella sandramycini* KACC 20249(T)**	97
Leucospermum conocarpodendron	Root	LcR50	***Kribbella aluminosa* HKI 0478(T) / *Kribbella karoonensis* DSM 17344(T)**	100.0	***Kribbella aluminosa* HKI 0480**	100
Leucospermum conocarpodendron	Root	LcR52	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	96.9	Uncultured bacterium	98
Leucospermum conocarpodendron	Root	LcR59	***Kribbella sancticallisti* BC633(T)**	98.4	***Kribbella solani* DSA1(T)**	98
Leucospermum conocarpodendron	Root	LcR60	***Kribbella aluminosa* HKI 0478(T)**	99.8	***Kribbella aluminosa* HKI 0480**	99
Leucospermum conocarpodendron	Root	LcR62	***Kribbella aluminosa* HKI 0478(T)**	99.8	***Kribbella aluminosa* HKI 0480**	99
Leucospermum conocarpodendron	Root	LcR63	***Kribbella amoyensis* XMU 198T**	99.4	***Kribbella flavida* DSM 17836(T)**	99
Leucospermum conocarpodendron	Root	LcR65	***Kribbella sancticallisti* BC633(T)**	99.2	***Kribbella solani* DSA1(T)**	99
Leucospermum conocarpodendron	Root	LcR69	Uncultured *Acidimicrobiales* GQ397046_s AK4DE1_07H	97.7	Uncultured actinobacterium	99
Orphium frutescens	Leaf	OfL1	Uncultured *Gemmatimonadales* EF632885_s Par-s-19	90.5	Uncultured bacterium	95
Orphium frutescens	Leaf	OfL2	***Kribbella solani* DSA1(T)**	99.1	***Kribbella solani* DSA1(T)**	99

Appendix 3C continued

Plant Name	Plant material	Clone Name	Ez-Taxon-e Top Hit	% Similarity	GenBank Top hit	% Similarity
Orphium frutescens	Leaf	OfL3	***Kribbella solani* DSA1 (T) / *Kribbella aluminosa* HKI 0478(T)**	99.3	***Kribbella solani* DSA1(T)**	99
Orphium frutescens	Leaf	OfL5	***Kribbella shirazensis* UTMC 693(T)**	99.2	***Kribbella aluminosa* HKI 0480**	99
Orphium frutescens	Leaf	OfL6	Uncultured *Acidimicrobiales* GQ397046_s AK4DE1_07H	97.1	Uncultured actinobacterium	98
Orphium frutescens	Leaf	OfL7	***Kribbella karoonensis* DSM 17344(T)**	82.5	***Kribbella solani* DSA1(T)**	82
Orphium frutescens	Leaf	OfL8	***Kribbella aluminosa* HKI 0478(T)**	99.8	***Kribbella aluminosa* HKI 0480**	99
Orphium frutescens	Leaf	OfL9	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	97.9	Uncultured bacterium	99
Orphium frutescens	Leaf	OfL12	***Kribbella solani* DSA1(T)**	94.5	***Kribbella flavida* DSM 17836(T)**	94
Orphium frutescens	Leaf	OfL16	Uncultured *Ilumatobacter* GU230444_s ARTE4_260	96.1	Uncultured bacterium	97
Orphium frutescens	Leaf	OfL19	Uncultured *Gemmatimonadales* EF516935_s FCPN548	96.9	Uncultured bacterium	97
Orphium frutescens	Leaf	OfL20	Uncultured *Nocardioides* FM873019_s FB02H10	50.7	Uncultured bacterium	95
Orphium frutescens	Leaf	OfL21	***Kribbella aluminosa* HKI 0478(T)**	99.8	***Kribbella aluminosa* HKI 0480**	99
Orphium frutescens	Leaf	OfL23	Uncultured *Acidimicrobiales* FJ203457_s SHFH508	52.5	Uncultured bacterium	90
Orphium frutescens	Leaf	OfL24	Uncultured *Acidimicrobiales* GQ397046_s AK4DE1_07H	97.5	Uncultured actinobacterium	98
Orphium frutescens	Leaf	OfL27	Uncultured *Gemmatimonadales* EF516935_s FCPN548	94.9	Uncultured bacterium	95
Orphium frutescens	Leaf	OfL29	Uncultured *Gaiella* JF429126_s C73	93.9	Uncultured actinobacterium	94
Orphium frutescens	Leaf	OfL31	***Kribbella aluminosa* HKI 0478(T)**	99.6	***Kribbella aluminosa* HKI 0480**	99
Orphium frutescens	Leaf	OfL32	Uncultured *Gemmatimonadales* AJ532728_s JG34-KF-418	97.1	Uncultured bacterium	96
Orphium frutescens	Leaf	OfL33	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	97.3	Uncultured bacterium	98
Orphium frutescens	Leaf	OfL38	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	97.5	Uncultured bacterium	98
Orphium frutescens	Leaf	OfL39	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	97.0	Uncultured bacterium	98
Orphium frutescens	Leaf	OfL41	Uncultured *Iamiaceae* FJ479385_s p22n13ok	49.2	Uncultured bacterium	99
Orphium frutescens	Leaf	OfL45	Uncultured *Gemmatimonadetes* EU134894_s FFCH6442	90.7	Uncultured bacterium	94
Orphium frutescens	Leaf	OfL46	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	96.6	Uncultured bacterium	97
Orphium frutescens	Leaf	OfL48	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	94.9	Uncultured bacterium	94
Orphium frutescens	Leaf	OfL63	Uncultured *Acidimicrobiales* GQ397046_s AK4DE1_07H	97.3	Uncultured actinobacterium	98
Orphium frutescens	Leaf	OfL64	***Kribbella aluminosa* HKI 0478(T)**	99.8	***Kribbella aluminosa* HKI 0480**	99
Orphium frutescens	Leaf	OfL65	***Kribbella sancticallisti* BC633(T)**	99.4	***Kribbella solani* DSA1(T)**	99
Orphium frutescens	Leaf	OfL66	Uncultured *Acidimicrobiales* GQ397046_s AK4DE1_07H	95.2	Uncultured actinobacterium	99
Orphium frutescens	Leaf	OfL68	***Kribbella sancticallisti* BC633(T)**	99.4	***Kribbella solani* DSA1(T)**	99
Orphium frutescens	Leaf	OfL69	***Kribbella sancticallisti* BC633(T)**	99.0	***Kribbella solani* DSA1(T)**	99

Appendix 3C continued

Plant Name	Plant material	Clone Name	Ez-Taxon-e Top Hit	% Similarity	GenBank Top hit	% Similarity
Orphium frutescens	Leaf	OfL72	***Kribbella aluminosa* HKI 0478(T)**	99.6	***Kribbella aluminosa* HKI 0480**	99
Orphium frutescens	Leaf	OfL73	***Kribbella aluminosa* HKI 0478(T)**	99.4	***Kribbella aluminosa* HKI 0480**	99
Orphium frutescens	Leaf	OfL74	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	97.9	Uncultured bacterium	98
Orphium frutescens	Leaf	OfL75	***Kribbella sancticallisti* BC633(T)**	97.7	***Kribbella solani* DSA1(T)**	97
Orphium frutescens	Leaf	OfL76	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	96.4	Uncultured bacterium	97
Orphium frutescens	Leaf	OfL77	Uncultured *Jatrophihabitans* AB245400_s Gsoil 1621	98.1	Actinomycetales bacterium Gsoil 1621	98
Orphium frutescens	Leaf	OfL78	***Kribbella solani* DSA1(T)**	98.7	***Kribbella solani* DSA1(T)**	99
Orphium frutescens	Leaf	OfL79	***Kribbella hippodromi* S1.4(T)**	99.4	***Kribbella solani* DSA1(T)**	99
Orphium frutescens	Root	OfR16	Uncultured *Acidimicrobiales* GQ397046_s AK4DE1_07H	97.5	Uncultured actinobacterium	98
Orphium frutescens	Root	OfR17	***Kribbella sandramycini* ATCC 39419(T)**	93.5	***Kribbella antibiotica* YIM31530(T)**	94
Orphium frutescens	Root	OfR19	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	97.9	Uncultured bacterium	98
Orphium frutescens	Root	*OfR21*	***Kribbella swartbergensis* HMC25(T)**	99.4	***Kribbella sandramycini* KACC 20249(T)**	99
Orphium frutescens	Root	OfR22	Uncultured *Lactobacillus* HM218836_s DMAR52	52.5	Cloning vector pG-Sable,	99
Orphium frutescens	Root	OfR23	Uncultured *Acidimicrobiales* AF448198_s ML817J-9	51.5	Uncultured actinobacterium	98
Orphium frutescens	Root	OfR24	*Blastococcus aggregatus* ATCC 25902(T)	97.0	***Kribbella sandramycini* KACC 20249(T)**	97
Orphium frutescens	Root	OfR26	Uncultured *Acidimicrobiales* GQ397046_s AK4DE1_07H	97.3	Uncultured actinobacterium	98
Orphium frutescens	Root	OfR30	***Kribbella sancticallisti* BC633(T)**	99.2	***Kribbella solani* DSA1(T)**	99
Orphium frutescens	Root	OfR31	***Kribbella sancticallisti* BC633(T)**	99.2	***Kribbella solani* DSA1(T)**	99
Orphium frutescens	Root	OfR33	***Kribbella aluminosa* HKI 0478(T)**	99.6	***Kribbella aluminosa* HKI 0480**	99
Orphium frutescens	Root	OfR34	Uncultured *lamiaceae* FM209118_s delph1F3	96.9	Uncultured bacterium	98
Orphium frutescens	Root	OfR35	***Kribbella sancticallisti* BC633(T)**	99.4	***Kribbella solani* DSA1(T)**	99
Orphium frutescens	Root	OfR36	***Kribbella aluminosa* HKI 0478(T) / *Kribbella karoonensis* DSM 17344(T)**	100.0	***Kribbella aluminosa* HKI 0480**	100
Orphium frutescens	Root	OfR39	***Kribbella solani* DSA1(T)**	99.6	***Kribbella solani* DSA1(T)**	99
Orphium frutescens	Root	OfR43	***Kribbella solani* DSA1(T)**	99.6	***Kribbella solani* DSA1(T)**	99
Orphium frutescens	Root	OfR45	***Kribbella sancticallisti* BC633(T)**	99.4	***Kribbella solani* DSA1(T)**	99
Orphium frutescens	Root	OfR48	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	96.2	Uncultured bacterium	96
Orphium frutescens	Root	OfR54	Uncultured *Acidimicrobiales* GQ397046_s AK4DE1_07H	97.5	Uncultured actinobacterium	98
Orphium frutescens	Root	OfR63	***Kribbella sancticallisti* BC633(T)**	98.0	***Kribbella flavida* DSM 17836(T)**	97
Orphium frutescens	Root	OfR64	***Kribbella aluminosa* HKI 0478(T) / *Kribbella karoonensis* DSM 17344(T)**	100.0	***Kribbella aluminosa* HKI 0480**	100
Orphium frutescens	Root	OfR65	***Kribbella aluminosa* HKI 0478(T)**	99.8	***Kribbella aluminosa* HKI 0480**	99

Appendix 3C continued

Plant Name	Plant material	Clone Name	Ez-Taxon-e Top Hit	% Similarity	GenBank Top hit	% Similarity
Orphium frutescens	Root	OfR66	***Kribbella aluminosa* HKI 0478(T)**	99.1	***Kribbella aluminosa* HKI 0480**	99
Orphium frutescens	Root	OfR69	***Kribbella solani* DSA1(T)**	99.4	***Kribbella solani* DSA1(T)**	99
Orphium frutescens	Root	OfR70	***Kribbella sancticallisti* BC633(T)**	99.2	***Kribbella solani* DSA1(T)**	99
Orphium frutescens	Root	OfR72	Uncultured *Jatrophihabitans* AB245400_s Gsoil 1621	98.1	*Actinomycetales* bacterium Gsoil 1621	98
Orphium frutescens	Root	OfR74	Uncultured *Acidimicrobiales* GQ397046_s AK4DE1_07H	97.7	Uncultured actinobacterium	99
Orphium frutescens	Root	OfR78	***Kribbella solani* DSA1(T)**	96.8	***Kribbella flavida* DSM 17836(T)**	96
Orphium frutescens	Root	OfR80	Uncultured *Iamiaceae* GQ487905_s T1-39 C73	49.0	Uncultured bacterium	98
Orphium frutescens	Root	OfR81	Uncultured *Gaiella* JF429126_s	94.5	Uncultured actinobacterium	95
Orphium frutescens	Root	OfR82	Uncultured *Ilumatobacter* EU799964_s 1C227642	48.7	Uncultured bacterium	99
Agathosma ovata 'glentana'	Leaf	AgL2	***Kribbella solani* DSA1(T)**	99.6	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL3	***Kribbella solani* DSA1(T)**	99.8	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL5	Uncultured *Paludibacter* AJ488070_s IA-16	49.9	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL6	***Kribbella hippodromi* S1.4(T)**	99.8	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL12	***Kribbella solani* DSA1(T)**	99.8	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL13	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	100.0	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL 15	***Kribbella solani* DSA1 (T) / *Kribbella aluminosa* HKI 0478(T)**	99.7	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL18	***Kribbella solani* DSA1(T)**	99.6	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL19	Uncultured *Methanomicrobia* AY714817_s Gzfos12E1	50.8	*Actinobacterium* SAUE53-11	99
Agathosma ovata 'glentana'	Leaf	AgL26	***Kribbella solani* DSA1(T)**	99.6	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL27	*Blastococcus aggregatus* ATCC 25902(T)	98.4	*Blastococcus aggregatus* DSM 4725(T)	98
Agathosma ovata 'glentana'	Leaf	AgL28	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	100.0	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL29	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	100.0	***Kribbella hippodromi* S1.4(T)**	100
Agathosma ovata 'glentana'	Leaf	AgL32	*Pseudomonas toyotomiensis* HT-3(T)	96.2	*Pseudomonas stutzeri* CW	99
Agathosma ovata 'glentana'	Leaf	AgL33	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	100.0	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL34	***Kribbella solani* DSA1(T)**	99.6	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL35	***Kribbella solani* DSA1(T)**	99.6	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL37	***Kribbella solani* DSA1(T)**	99.6	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL40	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	100.0	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL41	***Kribbella solani* DSA1(T)**	99.8	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL42	***Kribbella solani* DSA1(T)**	99.6	***Kribbella solani* DSA1(T)**	99

Appendix 3C continued

Plant Name	Plant material	Clone Name	Ez-Taxon-e Top Hit	% Similarity	GenBank Top hit	% Similarity
Agathosma ovata 'glentana'	Leaf	AgL44	***Kribbella solani* DSA1(T)**	99.6	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL51	***Kribbella solani* DSA1(T)**	99.6	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL52	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	100.0	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL53	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	100.0	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL54	***Kribbella solani* DSA1(T)**	99.4	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL55	***Kribbella karoonensis* DSM 17344(T)**	99.5	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL56	***Kribbella solani* DSA1(T)**	99.6	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL57	***Kribbella solani* DSA1(T)**	99.8	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL58	***Kribbella solani* DSA1(T)**	99.1	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL60	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	100.0	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL62	***Kribbella hippodromi* S1.4(T)**	100.0	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL63	***Kribbella solani* DSA1(T)**	99.8	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL64	***Kribbella solani* DSA1(T)**	99.3	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL65	***Kribbella solani* DSA1(T)**	99.4	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL66	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	100.0	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL73	***Kribbella solani* DSA1(T)**	99.4	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL74	***Kribbella solani* DSA1(T)**	99.8	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL75	***Kribbella solani* DSA1(T)**	99.4	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL76	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	100.0	***Kribbella albertanoniae* BC 640(T)**	98
Agathosma ovata 'glentana'	Leaf	AgL77	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	100.0	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL79	***Kribbella solani* DSA1(T)**	99.4	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL80	***Kribbella solani* DSA1(T)**	99.8	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL82	***Kribbella solani* DSA1(T)**	99.4	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL83	***Kribbella solani* DSA1(T)**	99.8	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL84	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	100.0	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL85	***Kribbella hippodromi* S1.4(T) / *Kribbella solani* DSA1(T)**	100.0	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Leaf	AgL87	***Kribbella solani* DSA1(T)**	99.6	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Stem	AgS7	***Kribbella solani* DSA1(T)**	99.6	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Stem	AgS9	*Blastococcus saxobsidens* BC448(T)	98.1	*Blastococcus saxobsidens* DSM 44509(T)	98
Agathosma ovata 'glentana'	Root	AgRK1	***Kribbella karoonensis* DSM 17344(T)**	100.0	***Kribbella aluminosa* HKI 0480**	99

Appendix 3C continued

Plant Name	Plant material	Clone Name	Ez-Taxon-e Top Hit	% Similarity	GenBank Top hit	% Similarity
Agathosma ovata 'glentana'	Root	AgRK7	***Kribbella aluminosa* HKI 0478(T)**	99.8	***Kribbella aluminosa* HKI 0480**	99
Agathosma ovata 'glentana'	Root	AgRK11	***Kribbella aluminosa* HKI 0478(T) / *Kribbella karoonensis* DSM 17344(T)**	100.0	***Kribbella aluminosa* HKI 0480**	100
Agathosma ovata 'glentana'	Root	AgRK13	***Kribbella karoonensis* DSM 17344(T)**	100.0	***Kribbella aluminosa* HKI 0480**	99
Agathosma ovata 'glentana'	Root	AgRK14	***Kribbella karoonensis* DSM 17344(T)**	100.0	***Kribbella aluminosa* HKI 0480**	99
Agathosma ovata 'glentana'	Root	AgRK17	***Kribbella aluminosa* HKI 0478(T) / *Kribbella karoonensis* DSM 17344(T)**	100.0	***Kribbella aluminosa* HKI 0480**	100
Agathosma ovata 'glentana'	Root	AgRK24	***Kribbella aluminosa* HKI 0478(T)**	99.6	***Kribbella aluminosa* HKI 0480**	99
Agathosma ovata 'glentana'	Root	AgRK26	***Kribbella aluminosa* HKI 0478(T)**	99.6	***Kribbella aluminosa* HKI 0480**	99
Agathosma ovata 'glentana'	Root	AgRK30	***Kribbella sancticallisti* BC633(T)**	99.0	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Root	AgRK31	***Kribbella shirazensis* UTMC 693(T)**	99.4	***Kribbella aluminosa* HKI 0480**	99
Agathosma ovata 'glentana'	Root	AgRK39	***Kribbella sancticallisti* BC633(T)**	99.2	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Root	AgRK40	***Kribbella karoonensis* DSM 17344(T)**	100.0	***Kribbella aluminosa* HKI 0480**	99
Agathosma ovata 'glentana'	Root	AgRK49	***Kribbella sancticallisti* BC633(T)**	99.4	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Root	AgRK50	***Kribbella aluminosa* HKI 0478(T) / *Kribbella karoonensis* DSM 17344(T)**	100.0	***Kribbella aluminosa* HKI 0480**	100
Agathosma ovata 'glentana'	Root	AgRK52	Uncultured *Lactobacillus* HM218836_s DMAR52	53.6	Uncultured bacterium	95
Agathosma ovata 'glentana'	Root	AgRK53	***Kribbella shirazensis* UTMC 693(T)**	99.4	***Kribbella aluminosa* XMU153**	99
Agathosma ovata 'glentana'	Root	AgRK67	***Kribbella shirazensis* UTMC 693(T)**	99.4	***Kribbella aluminosa* XMU153**	99
Agathosma ovata 'glentana'	Root	AgRK78	***Kribbella aluminosa* HKI 0478(T)**	99.7	***Kribbella aluminosa* XMU153**	99
Agathosma ovata 'glentana'	Soil	K2 Ag S	Uncultured *Cyanobacteria* GQ302549_s sw-xj33	92.0	Uncultured bacterium	93
Agathosma ovata 'glentana'	Soil	K4 Ag S	Uncultured *Acidimicrobiales* GQ397046_s AK4DE1_07H	97.7	Uncultured actinobacterium	99
Agathosma ovata 'glentana'	Soil	K5 Ag S	***Kribbella aluminosa* HKI 0478(T)**	99.6	***Kribbella aluminosa* HKI 0480**	99
Agathosma ovata 'glentana'	Soil	K6 Ag S	Uncultured *Gemmatimonadales* EF516935_s FCPN548	96.9	Uncultured bacterium	97
Agathosma ovata 'glentana'	Soil	K7 Ag S	Uncultured *Gemmatimonadales* HQ190412_s BR49	97.3	Uncultured bacterium	97
Agathosma ovata 'glentana'	Soil	K11 Ag S	Uncultured *Gemmatimonadales* GQ472385_s 3S1-12	96.8	Uncultured bacterium	97
Agathosma ovata 'glentana'	Soil	K15 Ag S	Uncultured *Acidimicrobiales* GQ397046_s AK4DE1_07H	97.5	Uncultured actinobacterium	98
Agathosma ovata 'glentana'	Soil	K16 Ag S	Uncultured *Gaiella* JF429126_s C73	94.5	Uncultured actinobacterium	95
Agathosma ovata 'glentana'	Soil	K17 AgS	Uncultured *Gaiella* FJ478994_s p7m05ok	48.0	Uncultured bacterium	75
Agathosma ovata 'glentana'	Soil	K19 AgS	Uncultured *Iamiaceae* FM209069_s delph1B1	97.6	Uncultured bacterium	99
Agathosma ovata 'glentana'	Soil	K21 AgS	Uncultured *Gemmatimonadales* EU803569_s 5C231158	97.5	Uncultured bacterium	98
Agathosma ovata 'glentana'	Soil	K23 AgS	Uncultured *Iamiaceae* FM209069_s delph1B1	97.6	Uncultured bacterium	99
Agathosma ovata 'glentana'	Soil	K24 AgS	Uncultured *Ilumatobacter* GU230444_s ARTE4_260	97.3	Uncultured bacterium	98

Appendix 3C continued

Plant Name	Plant material	Clone Name	Ez-Taxon-e Top Hit	% Similarity	GenBank Top hit	% Similarity
Agathosma ovata 'glentana'	Soil	K25 AgS	***Kribbella sancticallisti* BC633(T)**	99.4	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Soil	K28 AgS	Uncultured *Gemmatimonadales* EU979110_s g44-M-139	75.6	Uncultured bacterium	96
Agathosma ovata 'glentana'	Soil	K31 AgS	Uncultured *Gemmatimonadales* HQ114125_s V201-110	98.3	Uncultured bacterium	99
Agathosma ovata 'glentana'	Soil	K34 AgS	Uncultured *Ilumatobacter* GQ396861_s AK1AB2_05C	97.9	Uncultured bacterium	98
Agathosma ovata 'glentana'	Soil	K36 AgS	Uncultured *Iamiaceae*	97.9	Uncultured bacterium	98
Agathosma ovata 'glentana'	Soil	K40 AgS	***Kribbella sancticallisti* BC633(T)**	99.4	***Kribbella solani* DSA1(T)**	99
Agathosma ovata 'glentana'	Soil	K42 AgS	Uncultured *Gemmatimonadetes* FN667400_s PS2636	98.5	Uncultured compost bacterium	98
Tokai Forest	Soil	P2A1	Uncultured *Jatrophihabitans* AF498729_s Ellin347	97.8	*Frankiaceae* bacterium KVD-unk-16	98
Tokai Forest	Soil	P2A5	Uncultured *Sporichthyaceae* AM697007_s BF0001B033	96.2	Uncultured bacterium	97
Tokai Forest	Soil	P2A6	Uncultured *Jatrophihabitans* AB245400_s Gsoil 1621	95.6	Uncultured actinobacterium	99
Tokai Forest	Soil	P2A10	Uncultured *Jatrophihabitans* AF498729_s Ellin347	96.6	Uncultured bacterium	96
Tokai Forest	Soil	P2A11	Uncultured *Jatrophihabitans* AF498729_s Ellin347	96.6	Uncultured bacterium	96
Tokai Forest	Soil	P2A12	Uncultured *Jatrophihabitans* HQ674865_s MWM2-85	97.5	Uncultured *Frankineae* bacterium	97
Tokai Forest	Soil	P2A13	*Sporichthya brevicatena IFO 16195(T)*	95.4	Uncultured *Frankineae* bacterium	97
Tokai Forest	Soil	P2A14	Uncultured *Jatrophihabitans* AB245400_s Gsoil 1621	98.1	*Actinomycetales* bacterium Gsoil 1621	98
Tokai Forest	Soil	P2B1	Uncultured *Illumatobacter* EU803988_s 5C231666	50.7	Unidentified bacterium	98
Tokai Forest	Soil	P2B3	Uncultured *Frankiales* GQ850574_s d118	97.3	Uncultured actinobacterium	98
Tokai Forest	Soil	P2B5	Uncultured *Jatrophihabitans* HQ674865_s MWM2-85	97.5	Uncultured *Frankineae* bacterium	97
Tokai Forest	Soil	P2B8	Uncultured *Gemmatimonadaceae* AY395343_s EB1024	98.6	Uncultured bacterium	98
Tokai Forest	Soil	P2B10	*Actinoplanes tereljensis MN07-A0371(T)*	97.4	*Actinoplanes rectilineatus* IFO 13941(T)	97
Tokai Forest	Soil	P2B12	Uncultured *Jatrophihabitans* AF498729_s Ellin347	98.3	Uncultured bacterium	99
Tokai Forest	Soil	P2B15	Uncultured Bacteria EF018754_s Amb_16S_1214	97.8	Uncultured bacterium	98
Tokai Forest	Soil	P2B16	Uncultured *Jatrophihabitans* AB245400_s Gsoil 1621	95.6	Uncultured actinobacterium	99
Tokai Forest	Soil	P2B17	Uncultured *Jatrophihabitans* HQ674865_s MWM2-85	98.1	Uncultured bacterium	98
Tokai Forest	Soil	P2B20	*Sporichthya brevicatena IFO 16195(T)*	96.4	Uncultured bacterium	96
Tokai Forest	Soil	S2A3	Uncultured *Jatrophihabitans* AB245400_s Gsoil 1621	97.6	Uncultured bacterium	98
Tokai Forest	Soil	S2A4	***Kribbella amoyensis* XMU 198(T)**	99.2	***Kribbella sandramycini* KACC 20249(T)**	99
Tokai Forest	Soil	S2A7	Uncultured *Frankiaceae* AY234742_s	95.5	Uncultured *Frankineae* bacterium	96
Tokai Forest	Soil	S2A10	***Kribbella amoyensis* XMU 198(T)**	99.0	***Kribbella sandramycini* KACC 20249(T)**	99
Tokai Forest	Soil	S2B3	Uncultured *Jatrophihabitans* EU861905_s bacnit55	98.7	Uncultured bacterium	99

Appendix 3C continued

Plant Name	Plant material	Clone Name	Ez-Taxon-e Top Hit	% Similarity	GenBank Top hit	% Similarity
Tokai Forest	Soil	S2B6	Uncultured *Jatrophihabitans* HQ674865_s MWM2-85	98.7	Uncultured bacterium	99
Tokai Forest	Soil	S2B8	Uncultured *Jatrophihabitans* HQ674865_s MWM2-85	98.5	Uncultured *Frankineae* bacterium	98

Chapter 4:
Multilocus Sequence Analysis (MLSA) of the genus *Kribbella*

A portion of this chapter has been published: **Curtis, S.M. and Meyers, P.R. (2012).** Multilocus sequence analysis of the actinobacterial genus *Kribbella. Syst. Appl. Microbiol.* **35**, 441-446.

Chapter 4: Multilocus Sequence Analysis (MLSA) of the genus *Kribbella*

Chapter 4: Multilocus Sequence Analysis (MLSA) of the genus *Kribbella*

4.1 Summary

Multilocus sequence analysis (MLSA) was used to refine the phylogenetic analysis of the genus *Kribbella*, which currently contains 20 species with validly-published names. Sequences were obtained for the 16S rRNA, *gyrB, rpoB, recA, relA* and *atpD* genes for 20 type strains of the genus plus seven non-type strains and four laboratory strains that were determined to represent new species. A five-gene concatenated sequence of 4099 nt was used to examine the phylogenetic relationships between the species of the genus *Kribbella*. Using the concatenated sequence of the *gyrB, rpoB, recA, relA* and *atpD* genes, most *Kribbella* type strains can be distinguished by a genetic distance of >0.04. For the four *Kribbella* species for which multiple strains are available, the type strains share sequence similarities of 100% with the non-type strains of the same species for all genes except the 16S rRNA gene. Each single-gene tree had an overall topology similar to that of the concatenated sequence tree. The single-gene *relA* tree, used here for the first time in MLSA of actinobacteria, had good bootstrap support, comparable to the *rpoB* and *atpD* gene trees, which had topologies closest to that of the concatenated sequence tree. This illustrates that *relA* is a useful addition in MLSA studies of the genus *Kribbella*. By using MLSA with the *gyrB*, *rpoB, recA, relA* and *atpD* genes, an improved tree structure was obtained with increased bootstrap support, increased branch lengths and an overall higher resolution of the species in the genus. It is proposed that concatenated *gyrB-rpoB-recA-relA-atpD* gene sequences be used for examining the phylogenetic relationships of new strains of *Kribbella* and for determining the closest phylogenetic relatives to be used for taxonomic comparisons. Furthermore, the concatenated gene sequences can be used as a species delineation tool by utilizing a genetic distance threshold.

4.2 Introduction

Taxonomy allows for the organization of large amounts of knowledge about organisms and places them into meaningful groups in which predictions and hypotheses about the organisms can be made [Prescott *et al*., 2002]. Taxonomy is essential for accurate scientific communication and for the correct identification of microbes [Prescott *et al*., 2002]. Taxonomy is a fundamental discipline that provides a stable framework for classification, identification and sustains advances in a wide array of fields such as, clinical diagnostics, microbial ecology and biotechnology [Sutcliffe *et al*., 2012]. In order to improve and aid in the discovery of new rare actinobacteria and novel natural products, there must be an improvement in taxonomy and our understanding of microbial ecology [Kurtböke, 2012].

For the identification and classification of bacteria, microbiologists have relied heavily on 16S rRNA gene sequences for over 20 years [Rajendhran and Gunasekaran, 2010]. The 16S rRNA molecule, however, has a number of well known shortcomings when used for phylogenetic analysis [Santos and Ochman, 2004]. The 16S rRNA gene spans only a small portion of the genome, has a slow evolution rate and does not encode a protein, therefore any insertions or deletions can easily distort and complicate a phylogenetic tree [Santos and Ochman, 2004]. In addition, although it is useful in identifying strains to the genus level, the 16S rRNA gene does not provide sufficient resolution for differentiation between closely related species. [Santos and Ochman, 2004; Staley, 2006]. In order to overcome these problems, it has been shown in a number of actinobacterial studies that concatenating the sequences of several protein encoding gene fragments provides a more robust tree topology compared to a tree based solely on 16S rRNA gene sequences [Carro *et al*., 2012; Dalmasso *et al*., 2011; Gomila *et al*., 2008; Gou *et al*., 2008; Rong *et al*., 2009; Rong and Huang, 2010]. This multilocus sequence analysis (MLSA) approach has been shown to be a useful tool in determining a more accurate evolutionary history and for separation of species [Carro *et al*., 2012; Dalmasso *et al*., 2011; Rajendhran and Gunasekaran, 2010; Staley, 2006].

Multilocus sequence analysis is based on multilocus sequence typing, which was first introduced by Maiden *et al.* [1998] and was developed for typing pathogenic strains based on sequence data of multiple housekeeping genes. MLSA uses the concatenated sequence data for phylogenetic analysis and has the advantage of having higher phylogenetic resolution than the standard 16S rRNA gene approach without having to sequence entire genomes. Thus, MLSA has been referred to as an intermediate step between the use of 16S rRNA genes and whole genome sequencing of species [Cole *et al*., 2010; Kämpfer 2012; Maiden *et al*., 1998]. Despite the possibility that the genes chosen for MLSA could be susceptible to horizontal gene transfer and the fact that there is no consensus as to which genes should be used in MLSA studies, MLSA offers a considerable advantage to distinguish closely related species over the currently used DNA-DNA hybridization (DDH) techniques, which are laborious and time consuming [Rong *et al*., 2009; Santos and Ochman, 2004]. In addition to its species-resolving power, MLSA offers the additional important advantage over DDH methods of allowing for databases to be generated and shared. Nevertheless, it is standard practice for MLSA data to be compared with DDH values, as DNA relatedness values determined by DDH are currently the benchmark against which MLSA results are validated [Carro *et al*., 2012; Rong *et al*., 2009; Tindall *et al*., 2010].

The genus *Kribbella* belongs to the family *Nocardioidaceae* (suborder *Propionibacterineae*) [Parte, 2014; Park *et al*., 1999] and currently consists of 20 species with validly-published names [Parte, 2014]. The type species of the genus is *Kribbella flavida*. The genus was established in 1999 by Park *et al.* [1999] who reclassified two strains previously thought to belong to the genus *Nocardioides* [Sohn *et al*., 2003]. The genus *Kribbella* has been defined based solely on 16S rRNA gene phylogenetic analysis and unique chemotaxonomic markers and species within the genus share a high 16S rRNA gene sequence similarity. The genus *Kribbella* is phylogenetically related to the genera *Nocardioides* and *Aeromicrobium*, yet contains distinct chemotaxonomic characteristics. *Kribbella* strains have LL-diaminopimelic acid (LL-DAP) as the diagnostic diamino acid in the cell wall peptidoglycan [Park *et al*., 1999]. The major fatty acids in the genus *Kribbella* are anteiso-$C_{15:0}$, iso-$C_{16:0}$ and iso-$C_{15:0}$, whereas anteiso-$C_{15:0}$ is seldom detected in *Nocardioides* and *Aeromicrobium* [Park *et al*., 1999]. The predominant menaquinone in the genus *Kribbella* is MK-9(H_4), whereas the genus *Nocardioides* contains MK-8(H_4) as the major menaquinone [Park *et al*., 1999]. The

major phospholipids also distinguish *Kribbella* from the other genera and include phosphatidylcholine, diphosphatidylglycerol, phosphatidylglycerol and phosphatidylinositol [Park *et al*., 1999]. A further difference with the genus *Aeromicrobium* is that *Kribbella* has extensively branched vegetative mycelium, whereas *Aeromicrobium* cells grow as rods or cocci [Park *et al*., 1999]. The genomic DNA G+C content of *Kribbella* strains is 66-77% [Cui *et al*., 2010; Sohn *et al*., 2003; Everest *et al*., 2013].

Kribbella is an ideal candidate for MLSA as it is a relatively small genus, which allows for all the species to be analyzed and for phylogenetic relationships to be determined for the entire genus. It has been shown that the 16S rRNA gene sequence does not provide enough resolution in the phylogenetic analysis of *Kribbella* strains and cannot distinguish between closely related strains due to the high sequence similarity shared as seen, for example, with the type strains of *Kribbella solani* and *Kribbella hippodromi* [Curtis and Meyers, 2012]. Previous studies on *Kribbella* using the *gyrB* gene led to the generation of a more robust phylogenetic tree than that based on the 16S rRNA gene, suggesting that including more than one gene in phylogenetic studies of *Kribbella* could further improve the accuracy of the phylogenetic tree [Kirby *et al*., 2010].

Past studies have proposed a set of guidelines for choosing genes for MLSA and suggested that at least five universally conserved protein encoding genes, each longer than 900 bases are chosen. The genes must also be ubiquitous in all cellular life, yet be present only in a single copy in the genome and not be prone to horizontal gene transfer. The genes should also be well spaced on the genome [Adékambi *et al*., 2011; Santos and Ochman, 2004; Zeigler, 2003]. The choice of genes for this MLSA study of *Kribbella* was informed by these guidelines and by the distribution of the chosen genes on the *Kribbella flavida* DSM 17836^{T} genome. In addition to the 16S rRNA gene, five other genes were chosen: the *gyrB* gene (DNA gyrase B subunit), the *rpoB* gene (RNA polymerase beta subunit), the *recA* gene (recombinase A) and the *atpD* gene (ATP synthase F1 beta subunit). These four genes have proven useful in MLSA studies in other actinobacteria [Carro *et al*., 2012; Rong *et al*., 2009; Rong and Huang, 2010]. The fifth gene chosen for the study was the *relA* gene (encoding ppGpp synthetase) and was chosen for its important role in the response to nutrient stress [Chatterji and Ojha, 2001] and its location on the genome, being distant from the other genes that were chosen for

the MLSA analysis. This is the first time that *relA* has been used in the phylogenetic analysis of an actinobacterial genus.

The goal of this study was to refine the understanding of the phylogenetic relationships of the species in the genus *Kribbella* and to provide additional information towards a DNA sequence-based replacement for DDH in defining new *Kribbella* species. This is consistent with the phylogenomic species concept proposed by Staley [Staley, 2006; Staley, 2009]. By using MLSA with the *gyrB*, *rpoB, recA, relA* and *atpD* genes, an improved tree structure was obtained with increased bootstrap support and branch lengths and an overall higher resolution of the species in the genus. This study also proposes the use of a species delineation tool utilizing a genetic distance threshold.

4.3 Materials and Methods

4.3.1 Strains and DNA extraction

The 20 type strains of the species of the genus *Kribbella* with validly-published names at 31 August 2014, [Parte, 2014] plus seven non-type strains and four yet-to-be-described species, namely, '*Kribbella podocarpi*' strain YPL1, '*Kribbella capetownensis*' strain YM53 , '*Kribbella speibonae*' strain YM55 and '*Kribbella italica*' strain BC637 were included in this study. For genomic DNA extraction, strains were grown in 10 ml ISP2 [Shirling and Gottlieb, 1966] for 72 h at 30 °C with constant shaking, except for *Kribbella antibiotica* which was grown at room temperature (22°C). DNA was extracted using a modified version of the method of Wang *et al*. [1996], increasing the lysozyme concentration to 20mg/mL and the digestion period to overnight [Everest *et al*., 2011].

4.3.2 PCR primers and primer design for genes rpoB, recA, relA and atpD

In order to obtain the gene sequences, gene-specific PCR primers were designed using sequence alignments generated in MEGA 5.05 [Tamura *et al*., 2011]. The primers used in this study were designed from the available full genome sequences of closely related strains in the families *Nocardioidaceae* and *Propionibacteriaceae* (both suborder *Propionibacterineae*) and the families *Nocardiopsaceae* and *Streptosporangiaceae* (both

suborder *Streptosporangineae*). The strains used to design each primer set are listed in Appendix 4A. The primer sequences designed in this study are presented in Table 4.1 with their annealing temperatures and binding locations. The *K. flavida* DSM 17836^T genome is 7.58 Mb in size (CP001736; a total of 7086 genes) and all gene locations are in reference to this type-species genome. For the amplification of the *gyrB* gene, the primers used were those designed by Kirby *et al*. [2010], which produce a sequence 1135 nt in length. The *gyrB* gene is located near the 5' end of the *K. flavida* DSM 17836^T genome (locus tag Kfla_0010). The primers KrecA230F and KrecA815R amplify a 519 nt fragment of the *recA* gene (locus tag Kfla_3003). The KrelA586Fand KrelA1186R, and KrelA1091F and KrelA1658R PCR primers produce two overlapping fragments which, when combined, produce a sequence 1014 nt in length (*relA* gene locus tag Kfla_4398). The KatpD546F and KatpD1164R primers produce a 573 nt fragment (*atpD* gene locus tag Kfla_5046). The primers rpoBK1750F and rpoBK2692R amplify an 858 nt fragment (*rpoB* gene locus tag Kfla_6143). Primers rpoBK2242F and rpoBK2208R were used as internal primers for sequencing the *rpoB* gene fragment.

Table 4.1 PCR primers used to amplify the *Kribbella* genes *rpoB, recA, relA* and *atpD*

Primer name	Primer Sequence	Primer length (nt)	Binding position*	Annealing temperature (°C)
rpoBK1750F	5'- CBMTSATCCCVTTYCTSGAGCACG -3'	24	1750-1773	60-64
rpoBK2242F	5'- TCGAYGCBCGBGABACCAAGC -3'	21	2242-2262	60-64
rpoBK2208R	5'- GCRTCGACCTCGTRYTCYTCG -3'	21	2208-2229	60
rpoBK2692R	5'- CCSARCGGGTTSAGSAYGAYGTC -3'	23	2692-2714	60
KrecA230F	5'- TSGTSGAGRTCTAYGGHCCSGARTC -3'	25	230-254	60
KrecA815R	5'- TAVADRATGTCRAAYTCKGCCTGC -3'	24	815-838	60
KrelA586F	5'- GAYATYCGSGTBYTSGTSATCAAGC -3'	25	586-610	60
KrelA1091F	5'- TSCCSAAGTWCAACATGTACCAGTC - 3'	25	1091-1115	60
KrelA1186R	5'- GYCYTCCTTGTACTTCCAGTGCGC -3'	24	1186-1209	60
KrelA1658R	5'- AKCDSITCCTTGCCSYKVTCGATCG -3'	25	1658-1682	60
KatpD546F	5'- AAGACSGTGHTSATYCARGAGATG -3'	24	546-569	60
KatpD1164R	5'- ATVGCGATGATGTCCTGVARBTCCTTG - 3'	27	1164-1190	60

*Binding-position numbering is based on the full-length gene sequences in the *K. flavida* DSM 17836^T genome (CP001736).

4.3.3 PCR amplification and sequencing

PCR was carried out in 50 µL reaction volumes. Each reaction contained 2 mM $MgCl_2$, 0.5U Kapa *Taq* polymerase (Kapa Biosystems, Cape Town, South Africa), 150 µM of each dNTP, 0.5 µM of each primer and 1000 ng template DNA. PCR was performed using a Techne TC-512 thermal cycler. The PCR program used was: an initial denaturation at 96°C for 2 min, 30 cycles of denaturation at 96°C for 45 s, annealing at 60°C for 30 s and extension at 72°C for 45s, followed by a final extension at 72°C for 5 min. The PCR reaction for the *rpoB* gene required 1.0 µM of each primer and an annealing temperature ranging from 60-64 °C. To eliminate non-specific primer binding seen in some of the amplifications, the *rpoB* gene fragments from *Kribbella alba, Kribbella swartbergensis, Kribbella aluminosa* and *K. antibiotica* were amplified at 62°C, while those from *Kribbella catacumbae* DSM 19601^T and *K. catacumbae* BC632 were amplified at 64°C; all other *rpoB* amplifications were carried out at 60°C. The PCR products were electrophoresed on 0.8% (w/v) agarose gels containing 0.8 µg/ml ethidium bromide and visualised on a ChemiDoc^{TM} XRS+ Molecular Imager® (Bio-Rad). The amplified DNA was purified using an MSB® Spin PCRapace kit (Invitek, Germany). DNA sequencing was performed as a service by Macrogen Inc., Seoul, South Korea. Chromatograms were edited with Chromas version 2.01 (Technelysium Pty Ltd, Australia) and sequences were assembled in DNAMAN version 5.2.9 (Lynnon BioSoft). Aligned sequences were edited manually using MEGA 5.05 [Tamura *et al*., 2011]. The 16S rRNA and *gyrB* gene sequences were obtained from GenBank, with the exception of the *gyrB* gene sequences from *Kribbella albertanoniae* BC640^T, *Kribbella amoyensis* NRBC 107914^T, *K. catacumbae* BC628, *K. catacumbae* BC630, *K. catacumbae* BC632, "*K. capetownensis'* YM53^T, *Kribbella endophytica* DSM 17941^T, '*K. italica*' BC637^T, '*K. podocarpi'* YPL1^T, *Kribbella sancticallisti* BC634, *Kribbella shirazensis* DSM 45490^T and '*K. speibonae'* YM55^T which were obtained in this study.

4.3.4 Phylogenetic and sequence analysis

For phylogenetic analyses, sequences were aligned using MEGA version 5.05 and maximum-likelihood [Felsenstein, 1981], maximum-parsimony [Takahashi and Nei, 2000] and neighbour-joining [Saitou and Nei, 1987] trees were constructed. Concatenated gene sequences were generated by joining the gene fragments in the following order (5' → 3'): *gyrB* (1135 nt), *rpoB* (858 nt), *recA* (519 nt), *relA* (1014 nt) and

atpD (573 nt). The resulting *gyrB-rpoB-recA-relA-atpD* in-frame concatenated gene sequence consisted of 4099 nt. Sequence assembly was performed using DNAMAN version 5.2.9. Genetic distances were calculated in MEGA 5.05 using the Kimura 2-parameter model [Kimura, 1980]. A plot of genetic distances versus published DDH values was generated using SigmaPlot®, Version 10.0.1 (Systat Software Inc).

4.4 Results

4.4.1 *Gene amplification and 16S rRNA gene tree*

All primers proved successful in amplifying the correct gene for all strains used. Each gene sequence was subjected to nucleotide-nucleotide BLAST analysis against the GenBank database to confirm the identity of the gene. Table 4.2 lists the accession number for all genes and strains used in this study. Trees were constructed using the maximum-likelihood method [Felsenstein, 1981]. The 16S rRNA gene tree (Figure 4.1) showed two distinct groups (I and II), but had short branch lengths and there was low bootstrap support for most of the branches. With the exception of the association of the strains of *K. hippodromi* and *K. solani*, all high bootstrap values (>90%) were for clusters of strains belonging to the same species. The type strain of *K. antibiotica* formed the deepest branch in the 16S-rRNA gene tree.

4.4.2 *Phylogenetic analysis*

The *gyrB-rpoB-recA-relA-atpD* concatenated gene tree showed two conserved groups (A in green and B in blue) with very high bootstrap support (96% for group A and 100% for group B; Figure 4.2). Group A contained all the strains in 16S-rRNA gene tree group II plus the type strains of *K. antibiotica, K. albertanoniae*, *K. yunnanensis, K. alba, K. lupini, K. endophytica and K. flavida.* Group B contained most of the strains of 16S-rRNA gene tree group I, with the exceptions of *K. sandramycini, K. albertanoniae and K. yunnanensis*. The number of branches with greater than 90% bootstrap support increased from only four in the 16S rRNA gene tree (Figure 4.1) to 12 in the concatenated gene tree (Figure 4.2). Better resolution of *K. hippodromi* from the *K. solani* strains was also observed. Furthermore, the bootstrap support for the grouping of

Kribbella ginsengisoli and *Kribbella koreensis* increased from 84% to 100%. The concatenated-gene tree (Figure 4.2) had better resolution (i.e. longer branch lengths) and was more robust (i.e. higher bootstrap support) than the 16S rRNA gene tree (Figure 4.1) and the other single-gene trees (Appendices 4B-4F). There was also a marked improvement in the confidence one can place in the topology of the phylogenetic tree in going from the 16S rRNA gene tree (Figure 4.1) to the *gyrB* gene tree (Appendix 4B) to the *gyrB-rpoB-recA-relA-atpD* tree (Figure 4.2). The number of branches with bootstrap support greater than 90% increased from six in the *gyrB* gene tree to 12 in the *gyrB-rpoB-recA-relA-atpD* tree, demonstrating an improvement in the stability of the tree topology for the genus *Kribbella* on the addition of additional genes to the phylogenetic analysis.

Table 4.2. List of all *Kribbella* stains and accession numbers for gene sequences used in this study

Strain	16S rRNA gene accession numbers	*gyrB* gene accession numbers	*rpoB* gene accession numbers	*recA* gene accession numbers	*relA* gene accession numbers	*atpD* gene accession numbers
Kribbella alba DSM 15500^T	AY082062	EU434820	JX094181	JX094204	JX094227	JX094250
Kribbella albertanoniae BC640^T	KC283016	KC283007	KC283009	KC283011	KC283013	KC283015
Kribbella aluminosa DSM 18824^T	EF126967	EU434807	JX094182	JX094205	JX094228	JX094251
Kribbella aluminosa HKI 0479	EF126968	GQ244403	JX094183	JX094206	JX094229	JX094252
Kribbella aluminosa HKI 0480	EF1296969	GQ244404	JX094184	JX094207	JX094230	JX094253
Kribbella amoyensis NBRC 107914^T	HM368615	KC283006	KC283008	KC283010	KC283012	KC283014
Kribbella antibiotica DSM 15501^T	AY082063	EU434819	JX094185	JX094208	JX094231	JX094254
Kribbella catacumbae DSM 19601^T	AQUZ01000130	FJ917358	JX094186	JX094209	JX094232	JX094255
Kribbella catacumbae BC628	KM096441	JX094273	JX094187	JX094210	JX094233	JX094256
Kribbella catacumbae BC630	KM096442	JX094274	JX094188	JX094211	JX094234	JX094257
Kribbella catacumbae BC632	KJ875928	JX094275	JX094189	JX094212	JX094235	JX094258
Kribbella endophytica DSM 23718^T	HQ396152	KJ875918	KJ875920	KJ879522	KJ875924	KJ875926
Kribbella flavida CIP 107494^T	AY253863	NC_013729	JX094190	JX094213	JX094236	JX094259
Kribbella ginsengisoli DSM 17941^T	AB245391	JF775846	JX094191	JX094214	JX094237	JX094260
Kribbella hippodromi S1.4^T	EF472955	EU434817	JX094192	JX094215	JX094238	JX094261
'*Kribbella italica*' BC637^T	KJ875927	KJ875917	KJ875919	KJ875921	KJ875923	KJ875925
Kribbella jejuensis CIP 108509^T	AY253866	EU434818	JX094193	JX094216	JX094239	JX094262
Kribbella karoonensis Q41^T	AY995146	EU434816	JX094194	JX094217	JX094240	JX094263
Kribbella koreensis CIP 108301^T	Y09159	EU434810	JX094195	JX094218	JX094241	JX094264
Kribbella lupini LU14^T	AJ811962	EU434811	JX094196	JX094219	JX094242	JX094265
Kribbella sancticallisti DSM 19602^T	AM778577	FJ917357	JX094197	JX094220	JX094243	JX094266
Kribbella sancticallisti BC634	KM096443	JX094276	JX094198	JX094221	JX094244	JX094267
Kribbella sandramycini DSM 15626^T	AY253864	EU434812	JX094199	JX094222	JX094245	JX094268
Kribbella shirazensis DSM 45490^T	JN038072	KF908048	KF908049	KF908050	KF908051	KF908052
Kribbella solani CIP 108508^T	AY253862	EU434813	JX094200	JX094223	JX094246	JX094269
Kribbella solani YB2	EF623891	EU434814	JX094201	JX094224	JX094247	JX094270
Kribbella swartbergensis HMC25^T	AY995147	EU434808	JX094202	JX094225	JX094248	JX094271
Kribbella yunnanensis DSM 15499^T	AY082061	EU434815	JX094203	JX094226	JX094249	JX094272
'*Kribbella podocarpi' YPL1*	KM382222	KM382216	KM382213	KM382219	KM382210	KM382207
'*Kribbella capetownensis' YM53*	KM382223	KM382217	KM382214	KM382220	KM382211	KM382208
'*Kribbella speibonae' YM55*	KM382224	KM382218	KM382215	KM382221	KM382212	KM382209

* Accession numbers in blue text indicate sequences not obtained in this study

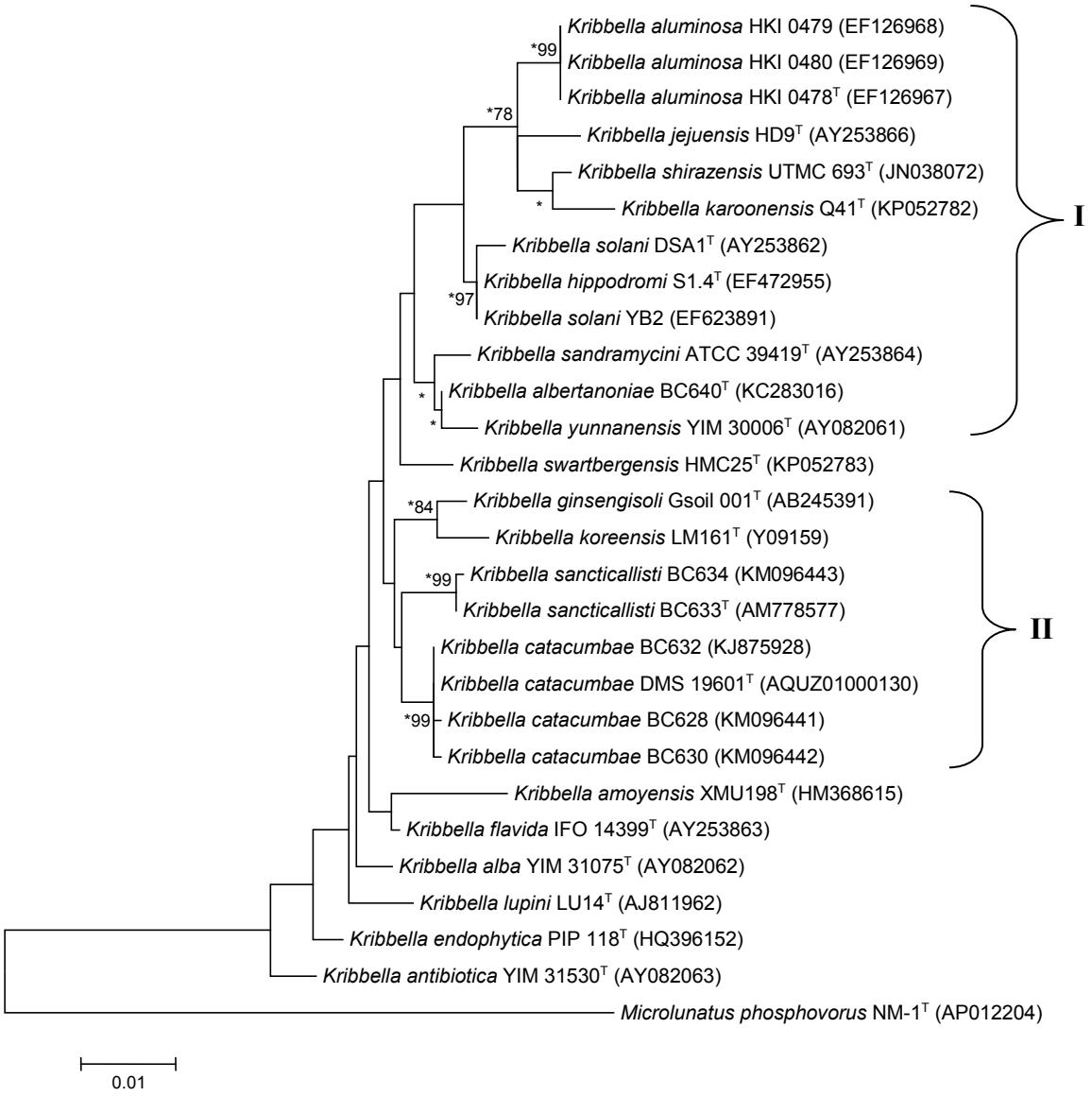


Figure 4.1 16S rRNA gene phylogenetic tree for the genus *Kribbella*. The tree is based on 1347 nt of common sequence and was constructed using the maximum-likelihood method. The bootstrap values are based on 1000 resampled datasets and only values greater than 70% are shown. Asterisks indicate the nodes that were also obtained in the maximum-parsimony and neighbour-joining trees. The scale bar represents 1 nt substitution per 100 nt. *Microlunatus phosphovorus* NM-1^T was used as an outgroup.

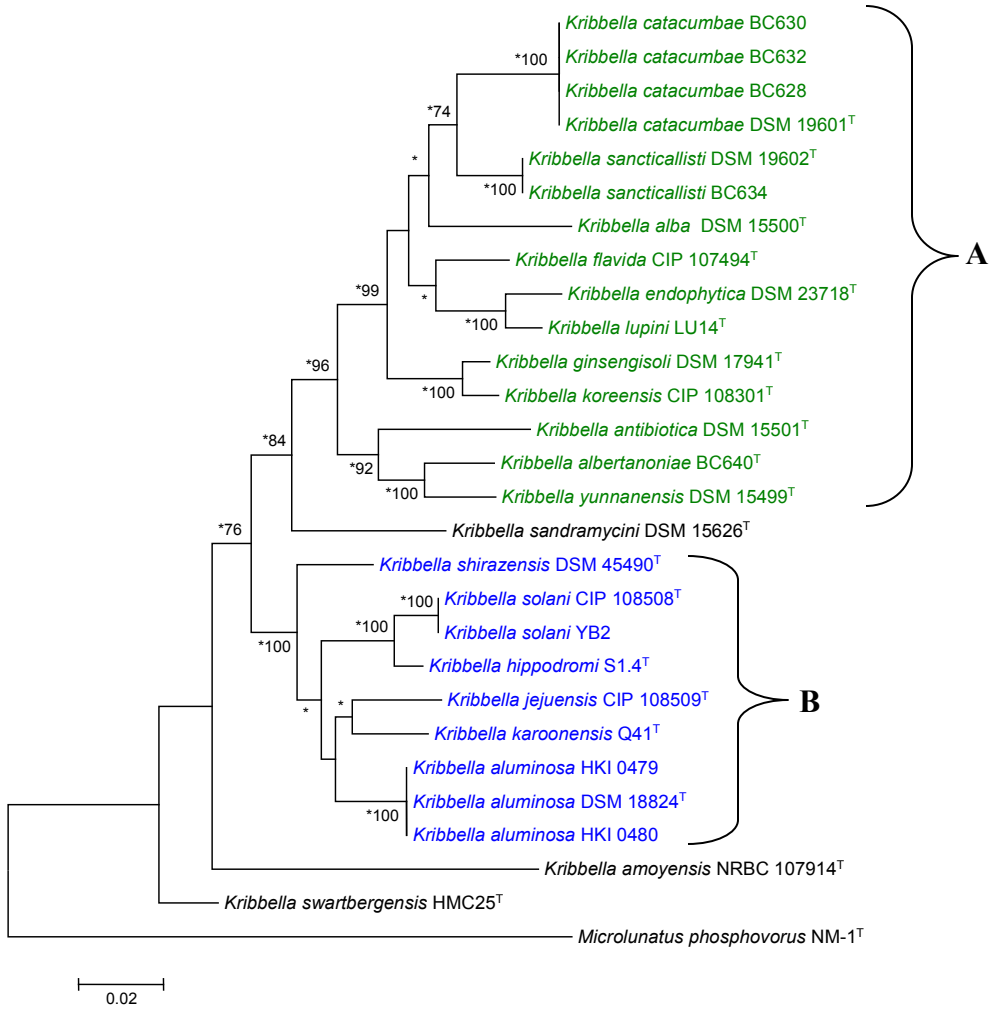


Figure 4.2. The concatenated *gyrB-rpoB-recA-relA-atpD* gene sequence phylogenetic tree for the genus *Kribbella*. The tree is based on 4099 nt of common sequence and was constructed using the maximum-likelihood method. The bootstrap values are based on 1000 resampled datasets and only values ≥ 70% are shown. Asterisks indicate the nodes that were also obtained in the maximum-parsimony and neighbour-joining trees. The scale bar represents 2 nt substitutions per 100 nt. *Microlunatus phosphovorus* NM-1^T was used as an outgroup.

A phylogenetic tree was also constructed after translation of the in-frame *gyrB-rpoB-recA-relA-atpD* concatenated gene sequences (Figure 4.3). Although the amino acid tree (Figure 4.3) showed the same two groups (A and B) as the tree based on concatenated nucleotide sequences (Figure 4.2), there was moderate bootstrap support for the association of the species of each group (84% for the strains of group A and 98% for the strains of group B) . The strains are colour coded to remain consistent with the groupings in Figure 4.2 to highlight the differences in Figure 4.3 for group A and B. *K amoyensis* NBRC 107914^T formed a long branch that was loosely associated with group

A and *K. sandramycini* DSM 15626^T was now loosely associated with group B. Overall, Figure 4.3 showed lower bootstrap values than the concatenated gene tree (Figure 4.2), with only five nodes ≥ 90% and, as expected, lower resolution (shorter branch lengths), because the amino acid sequences mask some of the underlying sequence differences seen at the nucleotide level due to the fact that some amino acids can be encoded by several codons. Nevertheless, the amino acid tree did resolve the species *K. solani* and *K. hippodromi*, which the 16S rRNA gene tree could not (Figure 4.1).

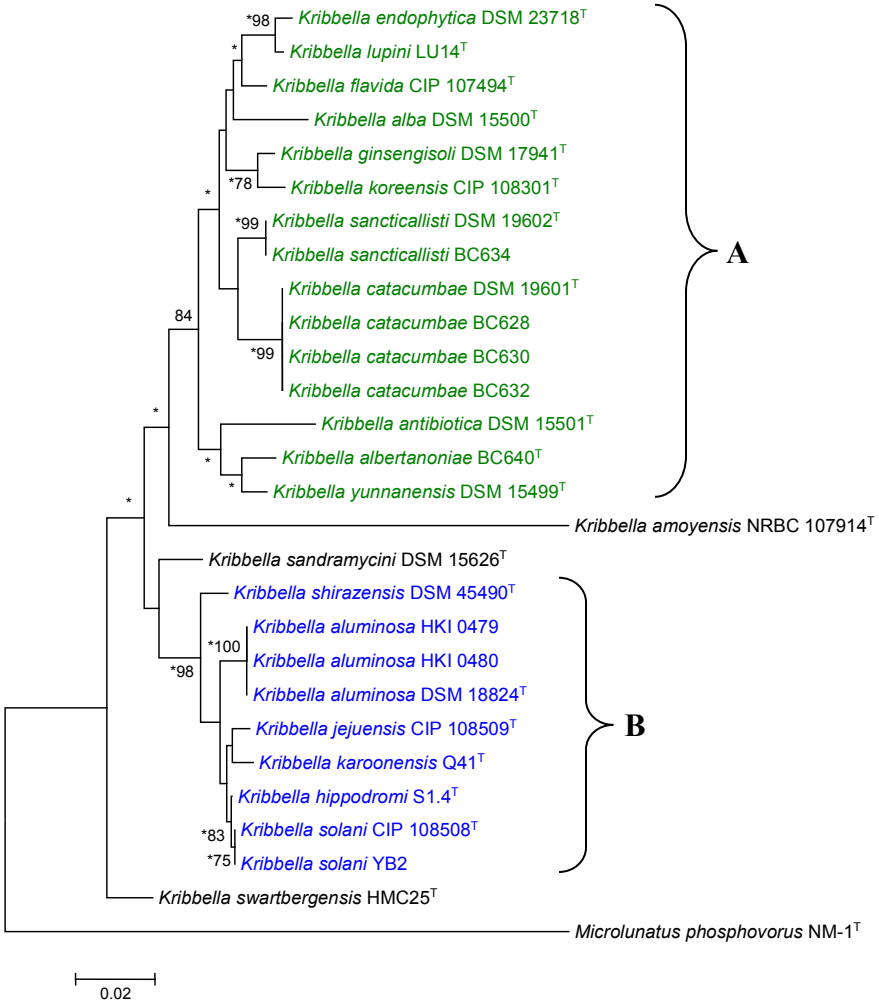


Figure 4.3 Phylogenetic tree for the genus *Kribbella* constructed from amino acid (aa) sequences obtained by translation of the concatenated, in-frame *gyrB-rpoB-recA-relA-atpD* gene sequences. The tree is based on 1342 aa of common sequence and was constructed using the maximum-likelihood method. The bootstrap values are based on 1000 resampled datasets and only values greater than 70% are shown. Asterisks indicate the nodes that were also obtained in the maximum-parsimony and neighbour-joining trees. The scale bar represents 2 aa substitutions per 100 aa. *Microlunatus phosphovorus* NM-1^T was used as an outgroup.

All of the single-gene trees for the protein-encoding genes (Appendices 4B-4F) grouped the strains into the two major clusters (designated A and B) seen in the *gyrB-rpoB-recA-relA-atpD* concatenated-gene sequence tree (Figure 4.2) with only minor changes. However, the *rpoB* tree (Appendix 4C) had the most similar topology to the *gyrB-rpoB-recA-relA-atpD* concatenated-gene sequence tree (Figure 4.2) with the sole exception of *K. sandramycini* being included in group A. The colour coding is based on Figure 4.2 with group A strains in green and group B strains in blue and is used to highlight differences in the single gene trees compared to the concatenated-gene sequence tree (Figure 4.2). The number of branches with bootstrap values of >90% were as follows: five (r*ecA*), six (*gyrB*), six (*atpD*), seven (*relA*) and eleven (*rpoB*).

As a gene to assess phylogenetic diversity, the *relA* gene provided a more robust tree (Appendix 4E) than that based on the 16S rRNA gene (Figure 4.1), with higher bootstrap values and longer branch lengths. However, in the *relA* gene tree, the group A species were split into three groups (albeit with weak bootstrap support) and the group A species *K. flavida* (Figure 4.2) was not closely associated with group A and formed the deepest branch in the tree (Appendix 4E).

4.4.3 *Genetic distances*

In addition to providing a more robust phylogenetic tree topology, the sequences of the five protein-encoding genes may also be used as a tool to assess overall sequence similarity between species. The DNA sequence similarities for all the type strains used in the study ranged from 97.1-99.6% (16S rRNA), 89.3-98.5% (*gyrB*), 91.5-99.2% (*rpoB*), 90.4-100% (*recA*), 90.3-99.1% (*relA*) and 92.0-99.8% (*atpD*). Interestingly, for the four *Kribbella* species for which multiple strains are available, the type strains share sequence similarities of 100% with the non-type strains of the same species for all genes except the 16S rRNA gene.

By plotting the genetic distances for the concatenated gene sequences (*gyrB-rpoB-recA-relA-atpD*) against published DDH values [Carlsohn *et al*., 2007; Cui *et al*., 2010; Curtis *et al*., manuscript under revision; Everest *et al*., 2015; Everest *et al*., 2013; Everest and Meyers, 2008; Kirby *et al*., 2006; Li *et al*., 2004; Li *et al*., 2006; Mohammadipanah *et al.,*

2013; Park *et al*., 1999; Sohn *et al*., 2003; Song *et al*., 2004; Trujillo *et al*., 2006; Urzì *et al*., 2008; Xu *et al*., 2012] (Figure 4.4), a threshold value can be determined such that a genetic distance value above the threshold corresponds to a value of less than 70% DNA-DNA relatedness, the threshold for distinguishing genomic species [Wayne *et al*., 1987]. From Figure 4.4, most *Kribbella* type strains are separated by concatenated-gene genetic distances of >0.04, with the exception of the comparisons of *K. hippodromi* with *K. solani* (concatenated-gene genetic distance = 0.018); *K. ginsengisoli* with *K. koreensis* (0.014) and *K. albertanoniae* with *K. yunnanensis* (0.033). The concatenated *gyrB-rpoB-recA-relA-atpD* genetic distance values are listed in Appendix 4H. Although having concatenated-gene genetic distances below 0.04, *K. hippodromi* S1.4^T has a DNA relatedness of only 40.4 % (± 3.8%) to *K. solani* DSA1^T [Everest and Meyers, 2008]; *K. ginsengisoli* DSM 17941^T has a DNA relatedness of 45% (± 5.37%) to *K. koreensis* LM161^T [Cui *et al*., 2010] and *K. albertanoniae* BC640^T has a DNA relatedness of 53 % (± 8%) to *K. yunnanensis* DSM 15499^T [Everest *et al*., 2013]. Therefore, for strains that are below the genetic distance threshold of 0.04, further testing (DDH) is recommended to determine if they are separate species.

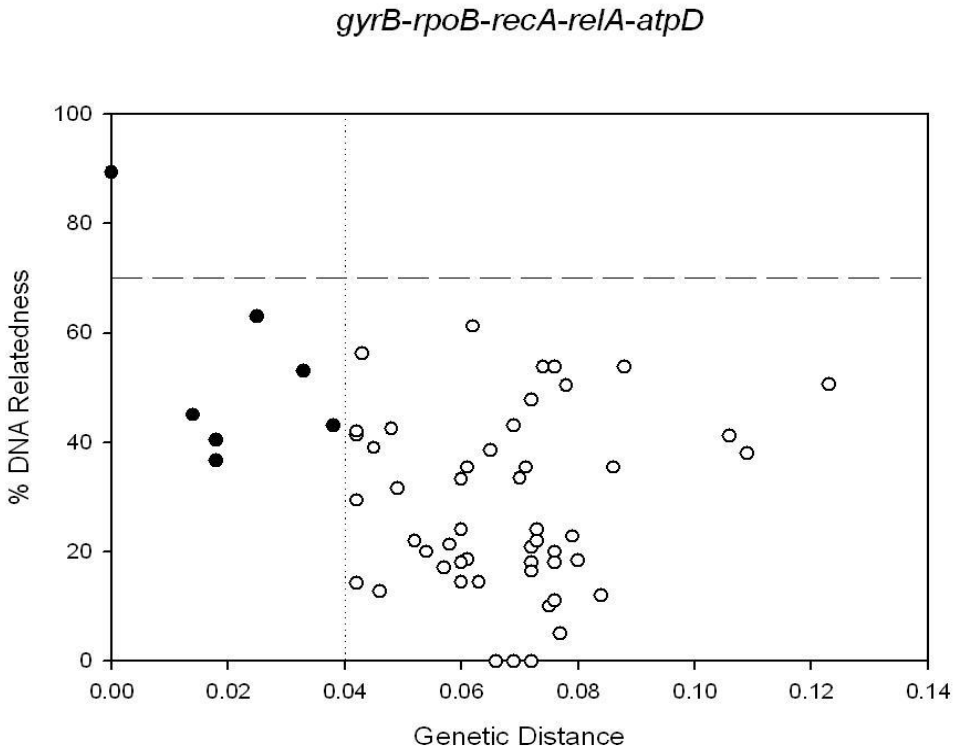


Figure 4.4. Graph of *gyrB-rpoB-recA-relA-atpD* based genetic distance versus DNA relatedness values for the genus *Kribbella*. The horizontal dashed line indicates the 70% DNA relatedness threshold, while the vertical dotted line indicates the proposed genetic distance threshold for distinguishing genomic species in the genus. The six points below the 0.04 threshold and under the 70% threshold represent seven pairwise comparisons of *K. hippodromi* S1.4^T vs. *K solani* CIP 108508^T (40.4% DNA relatedness by DDH) [Everest *et al*., 2011], *K hippodromi* S1.4^T vs. *K. solani* YB2, (36.6%) [Kirby *et al*., 2010], *K. ginsengisoli* DSM 17941^T vs. *K. koreensis* CIP 108301^T (45.0%) [Cui *et al*., 2010], *K. albertanoniae* BC640^T vs *K. yunnanensis* DSM 15499^T (53 %) [Everest *et al*., 2013], '*K. podocarpi*' YPL1^T vs *K. karoonensis* Q41^T (43%) [Curtis *et al*., manuscript under revision] and '*K. italica*' BC637^T vs *K. lupini* LU14^T and *K. endophytica* DSM 23718^T (both 63%) [Everest *et al*., 2015]. The comparison between *K. solani* CIP 108508^T and *K. solani* YB2 (89.3% DDH) [Kirby *et al*., 2010] is the only one under the 0.04 genetic distance threshold and above the 70% DNA relatedness threshold and shows that the strains belong to the same genomic species.

4.4.4 Genetic distance and MLSA for describing novel species

The approach of using genetic distances and MLSA was useful in the characterization of four previously isolated strains of *Kribbella* [Curtis *et al*., manuscript under revision;

Everest *et al*., 2015]. Firstly, in order to identify the strains, 16S rRNA and *gyrB* gene sequences were analysed. The 16S rRNA sequences were able to place the four isolates within the genus *Kribbella* with sequence similarities to *Kribbella* type strains ranging from 97.5% – 99.8% for strain YPL1, 97.5% – 99.0% for strain YM53, 97.7% – 99.3% for strain YM55 and 97.9% - 99.5% for strain BC637. The phylogenetic analysis utilizing 16S rRNA sequences (Figure 4.5) found that strain BC637 grouped with the type strain of *K. lupini* with low bootstrap support. Strain YM53 was loosely associated with the strains of group I and strains YM55 and YPL1 formed a group with the type strains of *K. karoonensis* and *K. shirazensis* within group I. In order to assess the isolates' closest phylogenetic relatives, the *gyrB*, *rpoB, recA, relA* and *atpD* genes were sequenced and incorporated into the MLSA for *Kribbella* (Figure 4.6). Using Figure 4.6, strain BC637 formed a very strongly supported group with the type strains of *K. lupini* and *K. endophytica* in group A. Strain YM53 formed a deep branch within the concatenated gene tree, as did the type strains of *K. swartbergensis* and *K. amoyensis*. Strains YM55 and YPL1 grouped with the strains of group B and strain YPL1 clustered with the type strain of *K. karoonensis* (with low bootstrap support). Strain YM55 formed the deepest branch in group B. Only minor changes were noted in the amino acid sequence tree (Figure 4.7) with regards to strains BC637, YM53, YM55 and YPL1. Strain YM53 formed a cluster with the type strains of *K. swartbergensis* and *K. sandramycini*, although with very low bootstrap support (Figure 4.7).

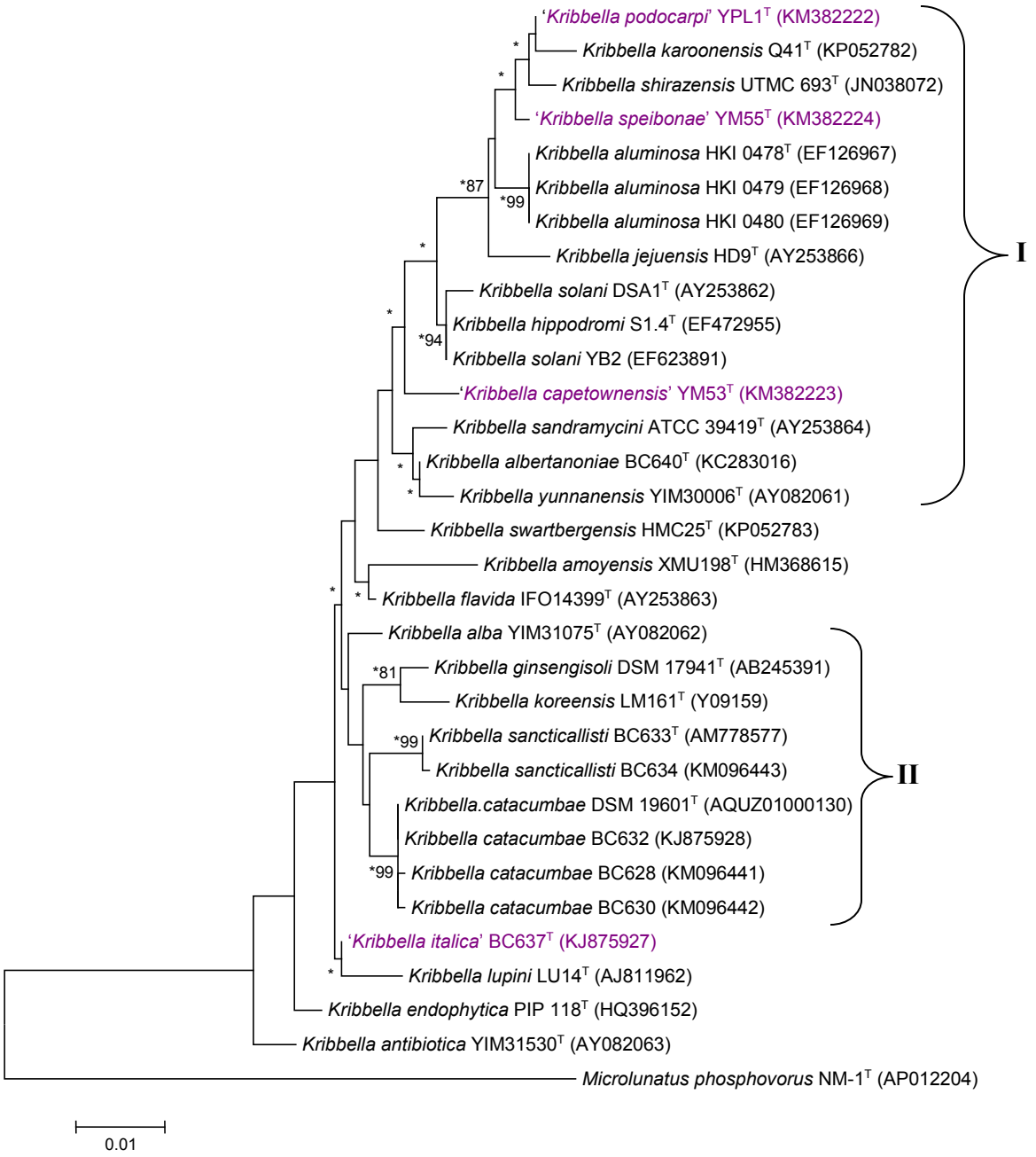


Figure 4.5. 16S rRNA gene phylogenetic tree for the genus *Kribbella* including novel strains highlighted in purple text. The tree is based on 1347 nt of common sequence and was constructed using the maximum-likelihood method. The bootstrap values are based on 1000 resampled datasets and only values greater than 70% are shown. Asterisks indicate the nodes that were also obtained in the maximum-parsimony and neighbour-joining trees. The scale bar represents 1 nt substitution per 100 nt. *Microlunatus phosphovorus* NM-1^T was used as an outgroup.

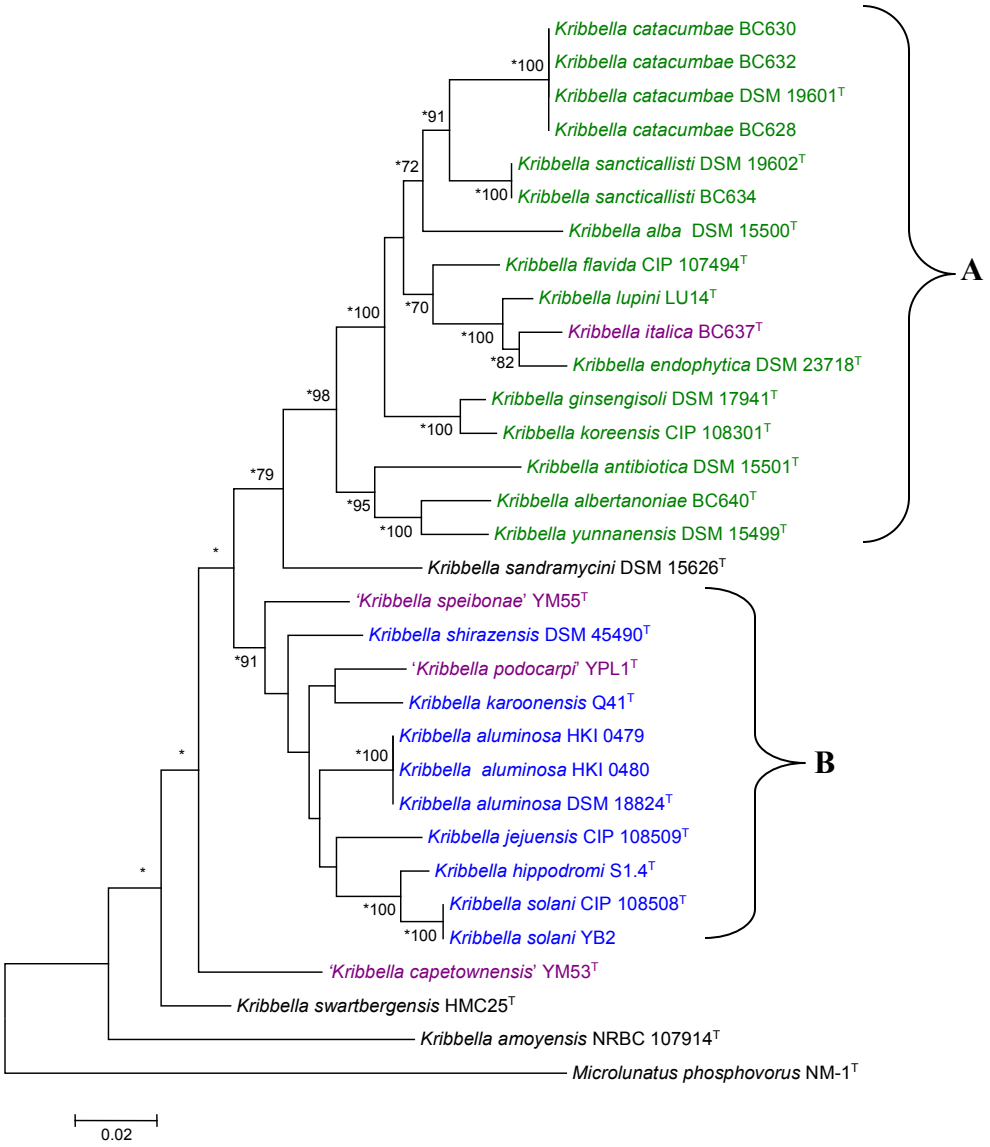


Figure 4.6. The concatenated *gyrB-rpoB-recA-relA-atpD* gene sequence phylogenetic tree for the genus *Kribbella* with the four novel isolates highlighted in purple text. The tree is based on 4099 nt of common sequence and was constructed using the maximum-likelihood method. The bootstrap values are based on 1000 resampled datasets and only values ≥ 70% are shown. Asterisks indicate the nodes that were also obtained in the maximum-parsimony and neighbour-joining trees. The scale bar represents 2 nt substitutions per 100 nt. *Microlunatus phosphovorus* NM-1^T was used as an outgroup.

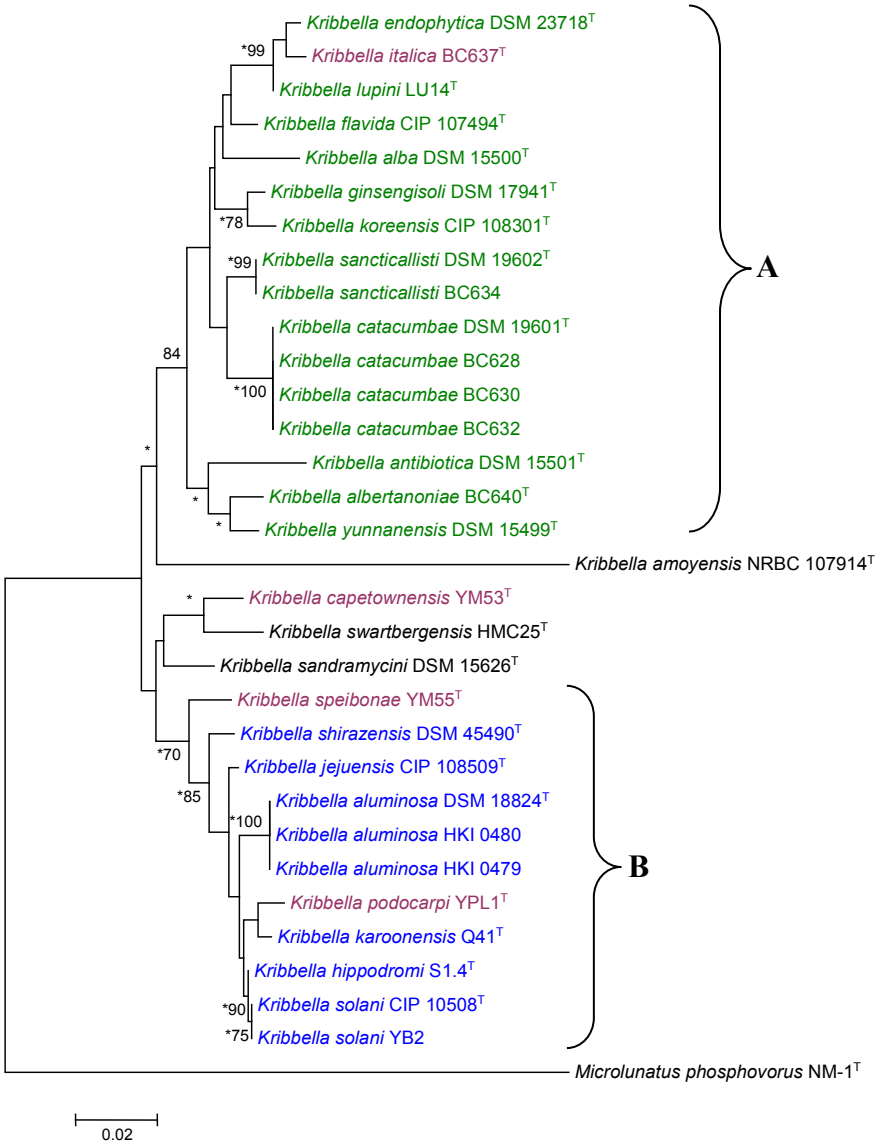


Figure 4.7. Phylogenetic tree for the genus *Kribbella* constructed from amino acid (aa) sequences obtained by translation of the concatenated, in-frame *gyrB-rpoB-recA-relA-atpD* gene sequences. The four novel isolates are highlighted in purple text. The tree is based on 1342 aa of common sequence and was constructed using the maximum-likelihood method. The bootstrap values are based on 1000 resampled datasets and only values ≥70% are shown. Asterisks indicate the nodes that were also obtained in the maximum-parsimony and neighbour-joining trees. The scale bar represents 2 aa substitutions per 100 aa. *Microlunatus phosphovorus* NM-1^T was used as an outgroup.

The *gyrB* sequences of the isolates were then compared against the *Kribbella* type strains using the method outlined by Kirby *et al.* [2010]. Strains YM55 and YM53 had *gyrB* genetic distance values against all *Kribbella* type strains that were above the 0.04

threshold. Strain YPL1, however, had one value under the 0.04 threshold (with the type strain of *K. karoonensis*; 0.032) and strain BC637 had two values under the threshold (with the type strains of *K. lupini* (0.026) and *K. endophytica* (0.016) [Curtis *et al*., manuscript under revision; Everest *et al*., 2015]). The concatenated-gene genetic distances (Figure 4.4) for the new isolates were similar to the *gyrB* genetic distance values, where strains YM53 and YM55 were above the threshold and strain YPL1 had one value under 0.04 (with the type strain of *K. karoonensis*; 0.038) and strain BC637 had two values under the threshold (with *K. lupini* (0.025) and *K .endophytica* (0.025)). As a result, DDH was performed between strain YPL1 and *K. karoonensis* Q41^T and strain BC637 was compared with *K. lupini* LU14^T and *K. endophytica* DSM 23718^T. It was found that strain YPL1 and *K. karoonensis* Q41^T had a 43% ± 13% DNA relatedness and, therefore, strain YPL1 represents a separate species [Curtis *et al.,* manuscript under revision]. The DNA relatedness value for strain BC637 and *K. lupini* LU14^T was 63% ± 0% and for strain BC637 and *K. endophytica* DSM 23718^T was 63% ±7%. Therefore, strain BC637 was determined to be a separate species [Everest *et al*., 2015]. DDH analyses were not undertaken using strains YM53 and YM55, as the genetic distance values between these strains and all *Kribbella* type strains were above the thresholds for both the *gyrB* and concatenated genes. From the concatenated gene tree (Figure 4.6), the closest phylogenetic neighbours of strains YPL1, YM53, YM55 and BC637 were established and further testing and physiological comparisons were then undertaken to fully characterize the strains for their description as the type strains of four novel species [Curtis *et al.,* manuscript under revision; Everest *et al*., 2015]. The species description paper for strains YPL1, YM53 and YM55 together and the species description paper for BC637 can be found in Appendix 4J and Appendix 4K, respectively.

4.5 Discussion

Genetic-distance thresholds can be utilized as a tool for determining if a new strain represents a novel species of *Kribbella*. It has been proposed that a *gyrB* gene genetic distance of ≥0.04 corresponds to a DNA relatedness by DDH of less than 70% for type

strains of the genus *Kribbella* [Kirby *et al*., 2010]. Using the concatenated gene sequences of *gyrB, rpoB, recA, relA* and *atpD*, a genetic distance of 0.04 is also proposed as a cut-off value for distinguishing type strains in this genus. Appendix 4G illustrates the use of *gyrB* gene genetic distances plotted against known DDH values (a table listing all genetic distance values for the *gyrB* gene are listed in Appendix 4I). Although there are pairs of strains with a *gyrB* genetic distance under the threshold of 0.04, there are fewer of these pairs than are seen in Figure 4.4 using the concatenated genes. The comparison of *K. albertanoniae* with *K. yunnanensis* had a *gyrB* genetic distance value of 0.052, while in the concatenated gene plot that value was below the threshold at 0.033. The *gyrB* gene genetic distance also placed the comparison of *K. ginsengisoli* and *K. koreensis* (0.042) above the threshold value, whereas the concatenated gene plot could not (0.018). Therefore, given the high species-resolving power of the *gyrB* gene in this genus, it would be quicker and simpler to compare *gyrB* genetic distances of new strains against all *Kribbella* type strains to determine the likelihood that the new strains represent new genomic species (before time and effort are invested in full taxonomic characterisation of the isolates). However, it would be beneficial for the concatenated *gyrB-rpoB-recA-relA-atpD* gene sequences to be utilized for exploring the phylogenetic relationships within the genus and for determining the closest phylogenetic relatives to be used for taxonomic comparisons.

The *rpoB* gene tree (Appendix 4C) on its own, had the most similar topology to that of the concatenated gene tree and thus the phylogenetic analysis from the single *rpoB* gene tree (Appendix 4C) would most closely resemble the full phylogenetic analysis from the concatenated gene tree. All the single gene trees were able to distinguish between closely related species, with the exception of the *recA* gene tree (Appendix 4D). The *recA* gene tree was able to distinguish between the closely related type strains of *K. hippodromi* and *K. solani*, which the 16S rRNA gene tree could not, but was unable to distinguish between the type strains of *K. ginsengisoli* and *K. koreensis*. This could indicate a possible horizontal gene transfer event of the *recA* gene between *K. ginsengisoli* and *K. koreensis*. Additionally, the *recA* analysis was based on only 519 nt of common sequence and therefore may not include enough data in order to separate the two species.

There are many examples of type strains that have a genetic distance under the threshold value for both *gyrB* and the concatenated gene genetic distance. Yet all the type strains tested had DNA similarity values under the 70% DDH threshold proposed by Wayne *et al.* [1987]. This suggests that both the *gyrB* and concatenated gene genetic distance thresholds could be lowered, since there are 4 pairs of type strains under the *gyrB* threshold (Appendix 4G) and 6 pairs of type strains under the concatenated threshold (Figure 4.4) that were proven to have DNA similarity values under the 70% cut-off.

By using an MLSA approach, the concatenated *gyrB-rpoB-recA-relA-atpD* gene sequences improved upon the phylogenetic analysis of the genus *Kribbella*. The 16S tree was unable to differentiate between *K. hippodromi* and *K. solani*, and had low bootstrap support and short branch lengths for the entire tree. The topology of the concatenated *gyrB-rpoB-recA-relA-atpD* gene tree was more robust with 12 branches with strong bootstrap support (>90%). This tree also had increased branch lengths, showing increased resolution of species. The concatenated *gyrB-rpoB-recA-relA-atpD* gene tree was also an improvement on the single *gyrB* gene tree. The 16S rRNA gene tree had low bootstrap values, implying low confidence in the topology of the tree and therefore in the relatedness of the clustered species. By increasing the number of genes in the analysis, the confidence in the topology of the tree increased (higher bootstrap values) and the resolution of the species also increased (longer branch lengths). This approach has also proved useful in refining the phylogenetic analysis of various actinobacterial genera, such as *Streptomyces, Micromonospora, Propionibacterium* and *Mycobacterium* [Carro *et al*., 2012; Dalmasso, *et al*., 2011; Gomila *et al*., 2008; Guo *et al*., 2008; Rong *et al*., 2009; Rong and Huang, 2010].

4.6 References

Adékambi, T., Butler, R. W., Hanrahan, F., Delcher, A. L., Drancourt, M. and Shinnick, T. M., (2011). Core gene set as the basis of multilocus sequence analysis of the subclass Actinobacteridae. *PLoS ONE*. **6,** e14792

Carlsohn, M.R., Groth, I., Spröer, C., Schütze, B., Saluz, H-P., Munder, T. and Stackebrandt E. (2007). *Kribbella aluminosa* sp. nov., isolated from a medieval alum slate mine. *Int. J. Syst. Evol. Microbiol*. **57**, 1943-1947.

Carro, L., Spröer, C., Alonso, P. and Trujillo, M.E. (2012). Diversity of *Micromonospora* strains isolated from nitrogen fixing nodules and rhizosphere of *Pisum sativum* analyzed by multilocus sequence analysis. *Syst. Appl. Microbiol*. **35**, 73-80.

Chatterji, D. and Ojha, A. K. (2001). Revisiting the stringent response, ppGpp and starvation signaling. *Current Opinion in Microbiology*. **4**, 160-165.

Cole, J.R., Konstantinidis, K., Farris, R.J., Tiedje, J.M. (2010). Microbial diversity phylogeny: extending from rRNAs to genomes. In: Liu, W.-T., Jackson, J.K. (eds). Environmental molecular microbiology. Caister Academic Press, Norfolk, 1-19.

Cui, Y.-S., Lee, J.-S., Lee, S.-T. and Im, W.-T. (2010). *Kribbella ginsengisoli* sp. nov., isolated from soil of a ginseng field. *Int. J. Syst. Evol. Microbiol*. **60**, 364-368.

Curtis, S.M. and Meyers, P.R. (2012). Multilocus sequence analysis of the actinobacterial genus *Kribbella. Syst. Appl. Microbiol.* **35**, 441-446.

Curtis, S.M., Norton, I., Everest, G.J. and Meyers, P.R. (2015). *Kribbella podocarpi* sp. nov., isolated from the leaves of a yellowwood tree (*Podocarpus latifolius*) and *Kribbella capetownensis* sp. nov. and *Kribbella speibonae* sp. nov., isolated from soil. *Antonie van Leeuwenhoek.* Manuscript under revision.

Dalmasso, M., Nicolas, P., Falentin, H., Valence, F., Tanskanen, J., Jatila, H., Salusjarvi, T. and Thierry A. (2011). Multilocus sequence typing of *Propionibacterium freudenreichii. Int. J. Food Microbiol*. **145**, 113-120.

Everest, G.J., Cook, A.E., Kirby, B.M. and Meyers, P.R., (2011). Evaluation of the use of *recN* sequence analysis in the phylogeny of the genus *Amycolatopsis. Antonie van Leeuwenhoek*. **100**, 483-496.

Everest, G.J., Curtis, S.M., De Leo, F., Urzí, C. and Meyers, P.R. (2013). *Kribbella albertanoniae* sp. nov., isolatated from a Roman catacomb, and emended description of the genus *Kribbella*. *Int. J. Syst. Evol. Microbiol*. **63**, 3591-3596.

Everest, G.J., Curtis, S.M., De Leo, F., Urzí, C. and Meyers, P.R. (2015). Description of *Kribbella italica* sp. nov., isolated from a Roman catacomb. *Int. J. Syst. Evol. Microbiol.* **64**, 000-000.*In Press*.

Everest, G.J. and Meyers, P.R., (2008). *Kribbella hippodromi* sp. nov., isolated from a soil from a racecourse in South Africa. *Int. J. Syst. Evol. Microbiol*. **58**, 443-446.

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *J. Mol. Evol*. **17**, 368-376.

Gomila, M., Ramirez, A., Gasco, J. and Lalucat. J. (2008). *Mycobacterium llatzerense* sp. nov., a facultatively autotrophic, hydrogen-oxidizing bacterium isolated from haemodialysis water. *Int. J. Syst. Evol. Microbiol*. **58**, 2769-2773.

Gou, Y., Zheng, W., Rong, X. and Huang, Y. (2008). A multilocus phylogeny of the *Streptomyces griseus* 16S rRNA gene clade: use of multilocus sequence analysis for streptomycete systematics. *Int. J. Syst. Evol. Microbiol.* **58**, 149-159.

Kämpfer, P. (2012). Systematics of prokaryotes: the state of the art. *Antonie van Leeuwenhoek.* **101**, 3-11.

Kirby, B.M., Everest, G.J. and Meyers, P.R., (2010). Phylogenetic analysis of the genus *Kribbella* based on the *gyrB* gene: proposal of a *gyrB*-sequence threshold for species delineation in the genus *Kribbella*. *Antonie van Leeuwenhoek*. **97**, 131-142.

Kirby, B.M., Le Roes, M. and Meyers, P.R. (2006). *Kribbella karoonensis* sp. nov. and *Kribbella swartbergensis* sp. nov., isolated from soil from the Western Cape, South Africa. *Int. J. Syst. Evol. Microbiol*. **56**, 1097-1101.

Kimura M. (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol*. **16**, 111-120.

Kurtböke, D.Í. (2012). Biodiscovery from rare actinomycetes: an eco-taxonomical perspective. *Appl. Microbiol. Biotechnol*. **93**, 1843-1852.

Li, W-J., Wang, D., Zhang, Y-Q., Schumann, P., Stackebrandt, E., Xu, L-H. and Jiang, C-L. (2004). *Kribbella antibiotica* sp. nov., a Novel Nocardioform Actinomycete Strain Isolated from Soil in Yunnan, China. *Syst. Appl. Microbiol*. **27**, 160-165.

Li, W.J., Wang, D., Zhang, Y.Q., Xu, L.H. and Jiang, C.L. (2006). *Kribbella yunnanensis* sp. nov., *Kribbella alba* sp. nov., two novel species of genus *Kribbella* isolated from soils in Yunnan, China. *Syst. Appl. Microbiol*. **29**, 29-35.

Maiden, M.C.J., Bygraves, J.A., Feil, E., Morelli, G., Russell, J.E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D.A., Feavers, I.M., Achtman, M. and Spratt, B.G. (1998). Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc. Natl. Acad. Sci. USA*. **95**, 3140-3145.

Mohammadipanah, F., Hamedi, J., Göker, M., Fiebig, A., Pukall, R., Spröer, C., Klenk, H.-P. (2013). *Kribbella shirazensis* sp. nov., isolated from Iranian soil. *Int. J. Syst. Bacteriol.* **63,** 3369-3374.

Park, Y.H., Yoon, J.H., Shine, Y.K., Suzuki, K., Kudo, T., Seino, A., Kim, H.J., Lee J.S. and Lee, S. T. (1999). Classification of '*Nocardioides fulvus'* IFO 14399 and *Nocardioides* sp. ATCC 39419 in *Kribbella* gen. nov., as *Kribbella flavida* sp. nov. and *Kribbella sandramycini* sp. nov. *Int. J. Syst. Bacteriol*. **49**, 743-752.

Parte, A. C. (2014). List of prokaryotic names with standing in nomenclature. (http://www.bacterio.net/kribbella.html) Accessed August 2014.

Prescott, L. M., Harley, J. P. & Klein, D. A. (2002). Microbial taxonomy. In: *Microbiology*, 5th Edition, pp. 421-449. USA: The McGraw-Hill Companies, Inc.

Rajendhran, J. and Gunasekaran, P., (2010). Microbial phylogeny and diversity: Small subunit ribosomal RNA sequence analysis and beyond. *Microbial Res*. **166**, 99-110.

Rong, X., Gou, Y. and Huang, Y., (2009). Proposal to reclassify the *Streptomyces albidoflavus* clade on the basis of multilocus sequence analysis and DNA-DNA hybridization, and taxonomic elucidation of *Streptomyces griseus* subsp. *solvifaciens*. *Syst. Appl. Microbiol*. **32**, 314-322.

Rong, X. and Huang, Y. (2010). Taxonomic evaluation of the *Streptomyces griseus* clade using multilocus sequence analysis and DNA-DNA hybridization, with proposal to combine 29 species and three subspecies as 11 genomic species. *Int. J. Syst. Evol. Microbiol*. **60**, 696-703.

Saitou, N. and Nei, M. (1987). The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol*. **4**, 406-425.

Santos, S.R. and Ochman, H. (2004). Identification and phylogenetic sorting of bacterial lineages with universally conserved genes and proteins. *Environ. Microbiol*. **6**, 754-759.

Sohn, K., Hong, S. G., Bae, K.S. and Chun, J. (2003). Transfer of *Hongia koreensis* Lee *et al*. 2000 to the genus *Kribbella* Park *et al*. 1999 as *Kribbella koreensis* comb. nov. *Int. J. Syst. Evol. Microbiol*. **53**, 1005-1007.

Song, J., Kim, B.Y., Hong, S.B., Cho, H.S., Sohn, K., Chun, J. and Suh, J.W. (2004). *Kribbella solani* sp. nov., and *Kribbella jejuensis* sp. nov., isolated from potato tuber and soil in Jeju, Korea. *Int. J. Syst. Evol. Microbiol*. **54**, 1345-1348.

Shirling, E.B. and Gottlieb, D., (1966). Methods for characterization of *Streptomyces* species. *Int. J. Syst. Bacteriol*. **16**, 313-340.

Staley, J.T. (2006). The bacterial species dilemma and the genomic-phylogenetic species concept. *Phil. Trans. R. Soc. B* **361**, 1899-1909.

Staley, J. (2009). The phylogenomic species concept. *Microbiol. Today* **36**, 80-83.

Sutcliffe, I.C., Trujillo, M.E. and Goodfellow, M. (2012). A call to arms for systematists: revitalising the purpose and practises underpinning the description of novel microbial taxa. *Antonie van Leeuwenhoek.* **101**, 13-20.

Takahashi, K. and Nei, M. (2000). Efficiencies of fast algorithms of phylogenetic inference under the criteria of maximum parsimony, minimum evolution and maximum likelihood when a large number of sequences are used. *Mol. Biol. Evol*. **17**, 1251-1258.

Tamura K., Peterson D., Peterson N., Stecher G., Nei M., and Kumar S. (2011). MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony. *Methods Mol. Biol. Evol.* **28**, 2731-2739.

Tindall, B.J., Rosselló-Móra, R., Busse, H.-J., Ludwig, W. and Kämpfer, P. (2010). Notes on the characterization of prokaryote strains for taxonomic purposes. *Int. J. Syst. Evol. Microbiol*. **60**, 249-266.

Trujillo, M.E., Kroppenstedt, R.M., Schumann, P. and Martínez-Molina, E. (2006). *Kribbella lupini* sp. nov., isolated from the roots of *Lupinus angustifolius*. *Int. J. Syst. Evol. Microbiol*. **56**, 407-411.

Urzì, C., De Leo, F. and Schumann, P. (2008). *Kribbella catacumbae* sp. nov. and *Kribbella sancticallisti* sp. nov., isolated from whitish-grey patinas in the catacombs of St. Callistus in Rome, Italy. *Int. J. Syst. Evol. Microbiol*. **58**, 2090-2097.

Wang, Y., Zhang, Z. and Ruan, J. (1996). A proposal to transfer *Microbispora bispora* (Lechevalier 1965) to a new genus, *Thermobispora* gen. nov., as *Thermobispora bispora* comb. nov. *Int. J. Syst. Bacteriol*. **46**, 933-938.

Wayne, L.G., Brenner, D.J., Colwell, R.R., Grimont, P.A.D., Kandler, O., Krichevsky, M.I., Moore, L.H., Moore, W.E.C., Murray, R.G.E., Stackebrandt, E., Starr, M.P. and Trüper, H.G. (1987). International Committee on Systematic Bacteriology. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *Int. J. Syst. Bacteriol*. **37**, 463-464.

Xu, Z., Xu, Q., Zheng, Z. and Huang, Y. (2012). *Kribbella amoyensis* sp. nov., isolated from rhizosphere soil of a pharmaceutical plant, *Typhonium giganteum* Engl. *Int. J. Syst. Evol. Microbiol*. **62**, 1081-1085.

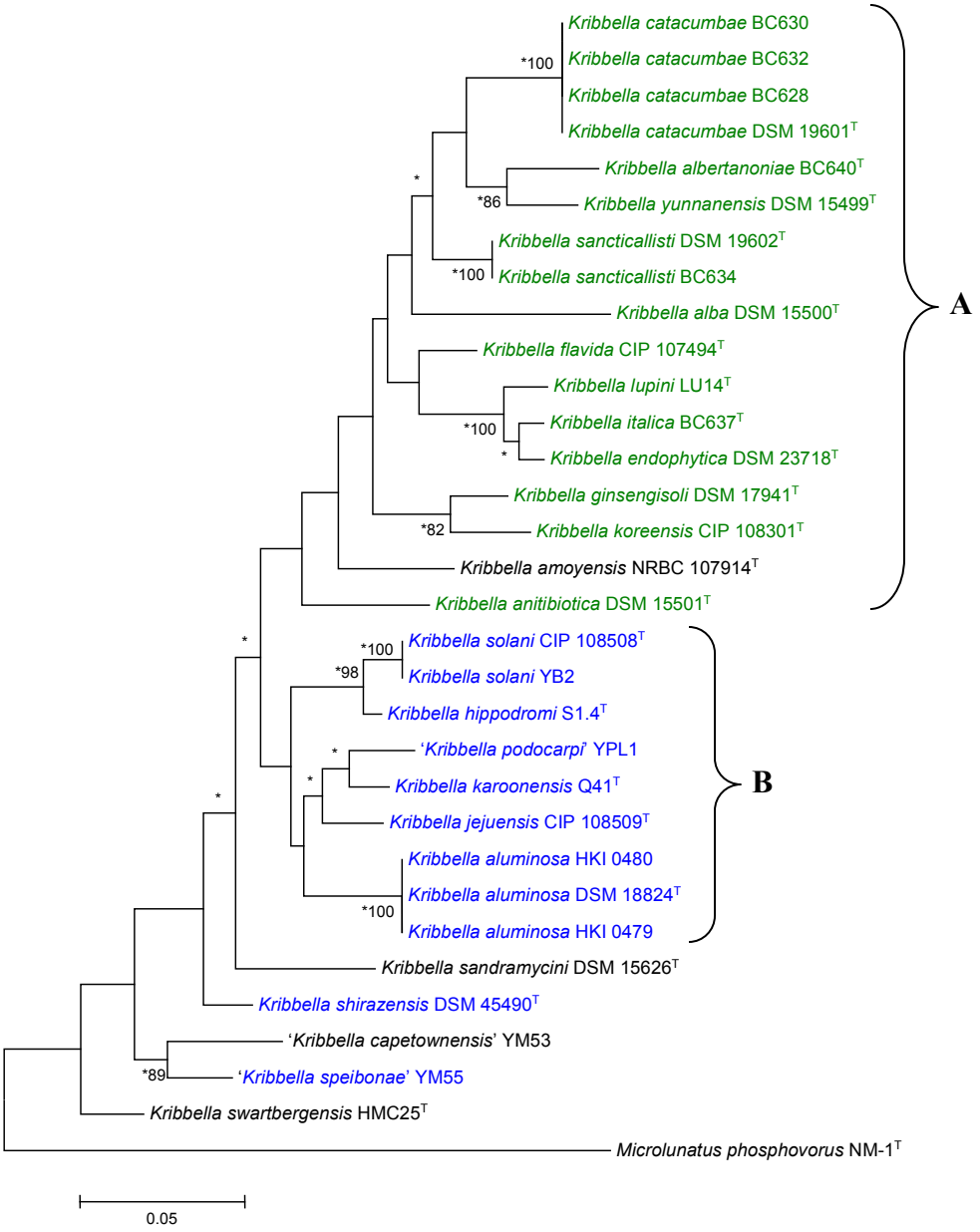
Zeigler, D. R. (2003). Gene sequences useful for predicting relatedness of whole genomes in bacteria. *Int. J. Syst. Evol. Microbiol*. **53**, 1893-1900.

4.7 Appendices

Appendix 4A: Strains used to design PCR primers

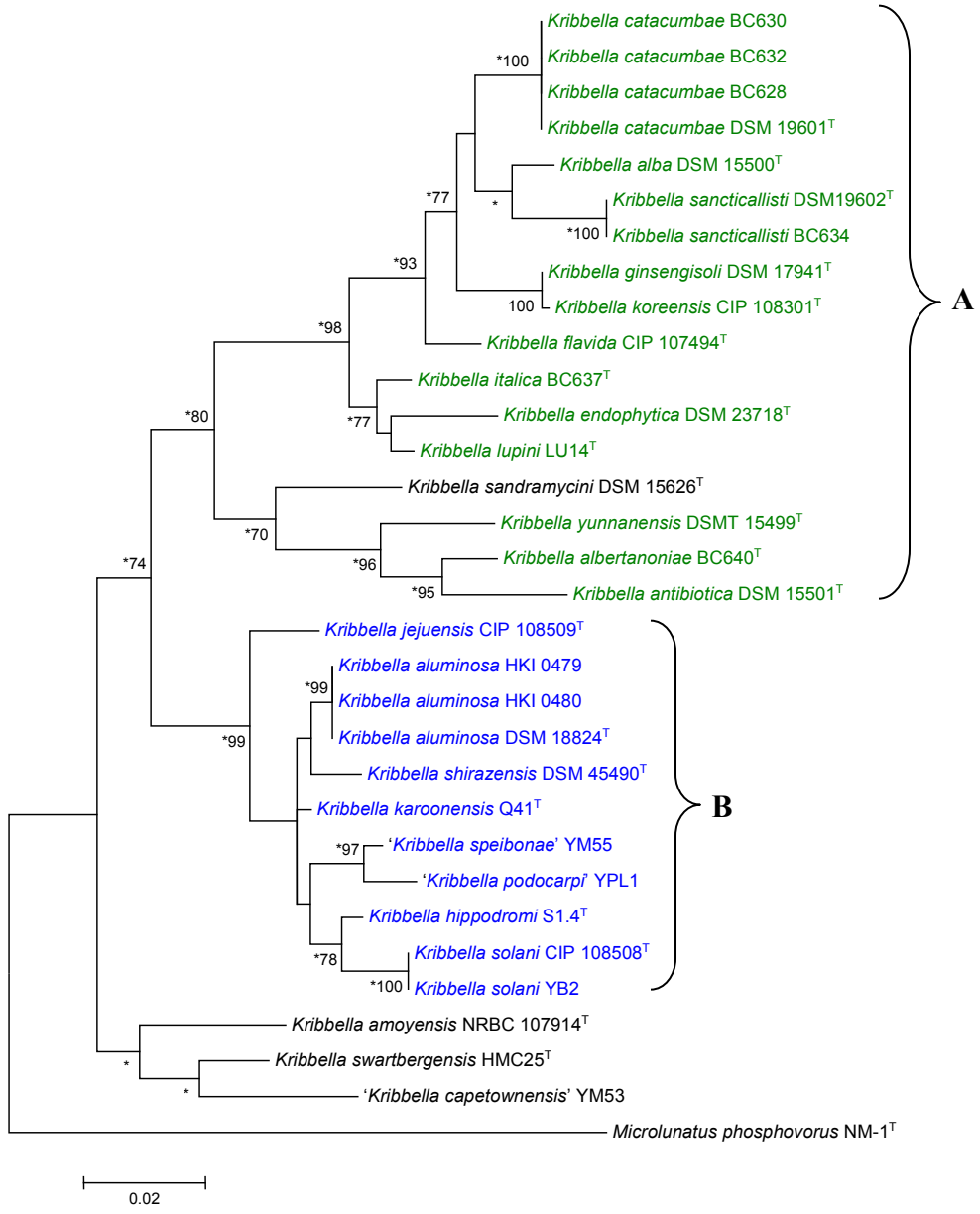
Primer	Strains Used
rpoBK1750F, rpoBK2242F, rpoBK2208R and rpoBK2696R	*Kribbella flavida* DSM 17836^T *Nocardioides* sp. JS614 *Propionibacterium freudenreichii* subsp. *shermanii* CIRM-BIA1^T *Propionibacterium acnes* KPA171202 *Propionibacterium acnes* SK137 *Nocardiopsis dassonvillei* subsp. *dassonvillei* DSM 43111^T *Streptosporangium roseum* DSM 43021^T *Thermobifida fusca* YX *Thermobispora bispora* DSM 43833^T *Thermomonospora curvata* DSM 43183^T
KrecA230F and KrecA815R	*Kribbella flavida* DSM 17836^T *Nocardioides* sp. JS614 *Nocardiopsis dassonvillei* subsp. *dassonvillei* DSM 43111^T, *Propionibacterium freudenreichii subsp. shermanii CIRM-BIA1T,* *Propionibacterium acnes KPA171202* *Propionibacterium acnes* SK137 *Streptosporangium roseum* DSM 43021^T *Thermobispora bispora* DSM 43833^T
KrelA586F, KrelA1091F, KrelA1186R	*Kribbella flavida* DSM 17836^T *Microlunatus phosphovorus* NM-1^T *Aeromicrobium marinum* DSM 15272^T *Nocardioides* sp. JS614 *Nocardiopsis dassonvillei* subsp. *dassonvillei* DSM 43111^T, *Propionibacterium freudenreichii subsp. shermanii CIRM-BIA1T* *Propionibacterium acnes KPA171202* *Propionibacterium acnes* SK137 *Propionibacterium acnes* 266 *Propionibacterium acnes* 6609
KrelA1658R	*Kribbella flavida* DSM 17836^T *Microlunatus phosphovorus* NM-1^T *Aeromicrobium marinum* DSM 15272^T, *Nocardioides* sp. JS614 *Propionibacterium freudenreichii* subsp. *shermanii* CIRM-BIA1^T
KatpD546F and KatpD1164R	*Kribbella flavida* DSM 17836^T *Microlunatus phosphovorus* NM-1^T *Aeromicrobium marinum* DSM 15272^T *Nocardioides sp*. JS614 *Propionibacterium freudenreichii subsp. shermanii* CIRM-BIA1^T *Propionibacterium acnes* KPA171202

Appendix 4B (*gyrB* tree)



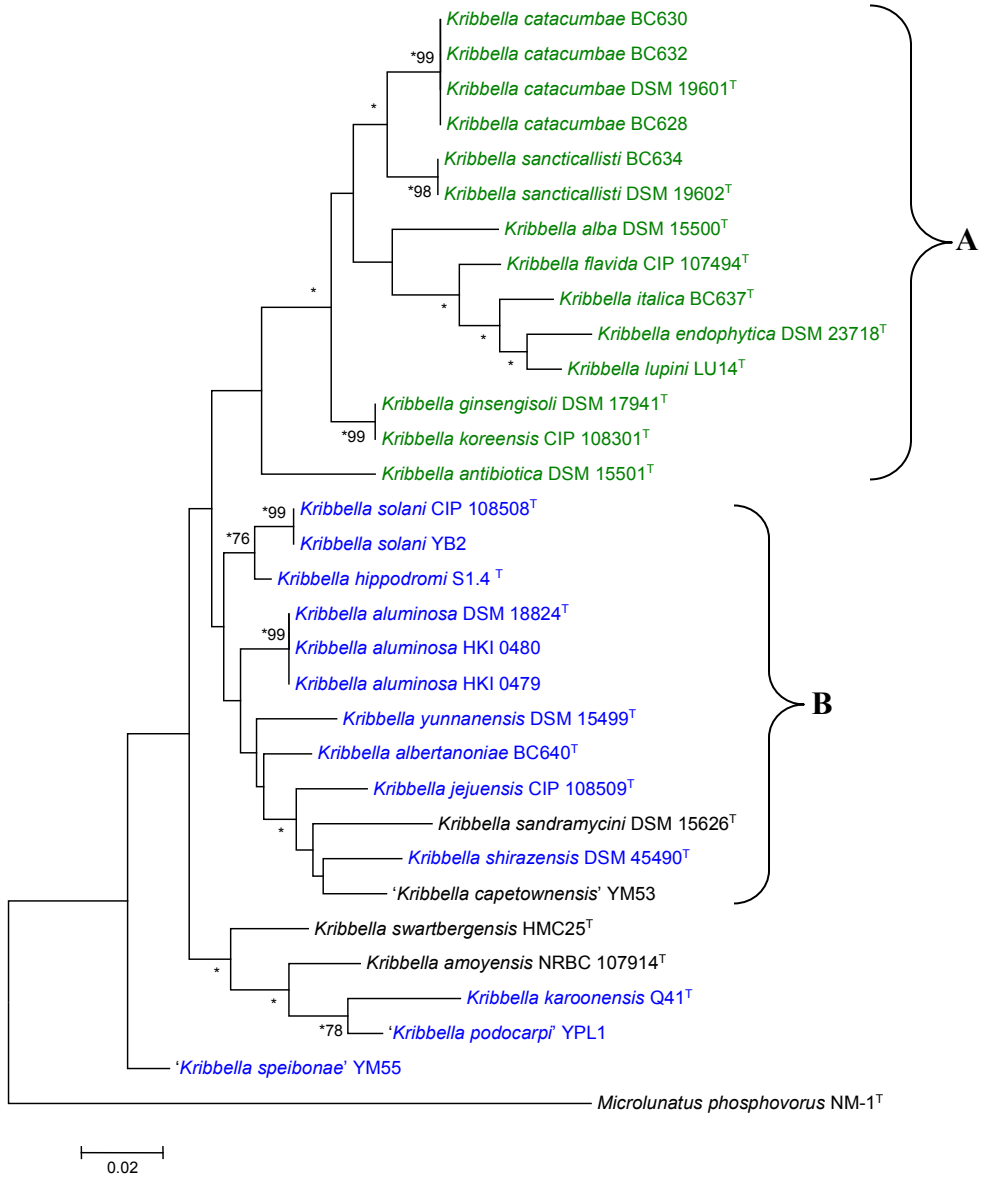
Appendix 4B. The *gyrB* gene sequence phylogenetic tree for the genus *Kribbella*. The tree is based on 1135 nt of common sequence and was constructed using the maximum-likelihood method. The bootstrap values are based on 1000 resampled datasets and only values greater than 70% are shown. Asterisks indicate the nodes that were also obtained in the maximum-parsimony and neighbour-joining trees. The scale bar represents 5 nt substitutions per 100 nt. *Microlunatus phosphovorus* NM-1^T was used as an outgroup.

Appendix 4C (*rpoB* tree)



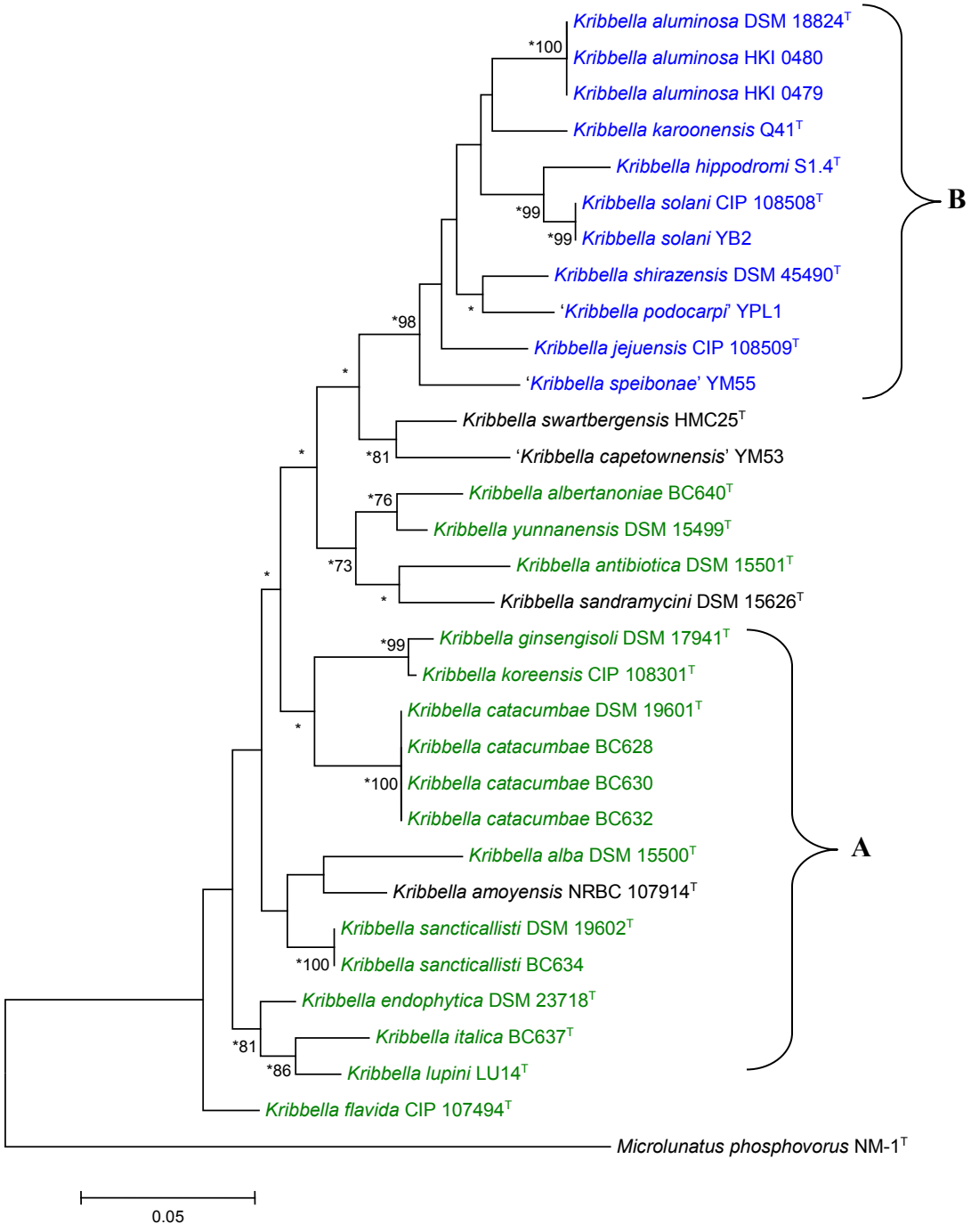
Appendix 4C: *rpoB* gene phylogenetic tree for the genus *Kribbella*. The tree is based on 858 nt of common sequence and was constructed using the maximum likelihood method. The bootstrap values are based on 1000 resampled datasets and only values ≥70% are shown. Asterisks indicate the nodes that were also obtained in the maximum-parsimony and neighbour-joining trees. The scale bar represents 2 nt substitutions per 100 nt. *Microlunatus phosphovorus* NM-1T was used as an outgroup.

Appendix 4D (*recA* tree)



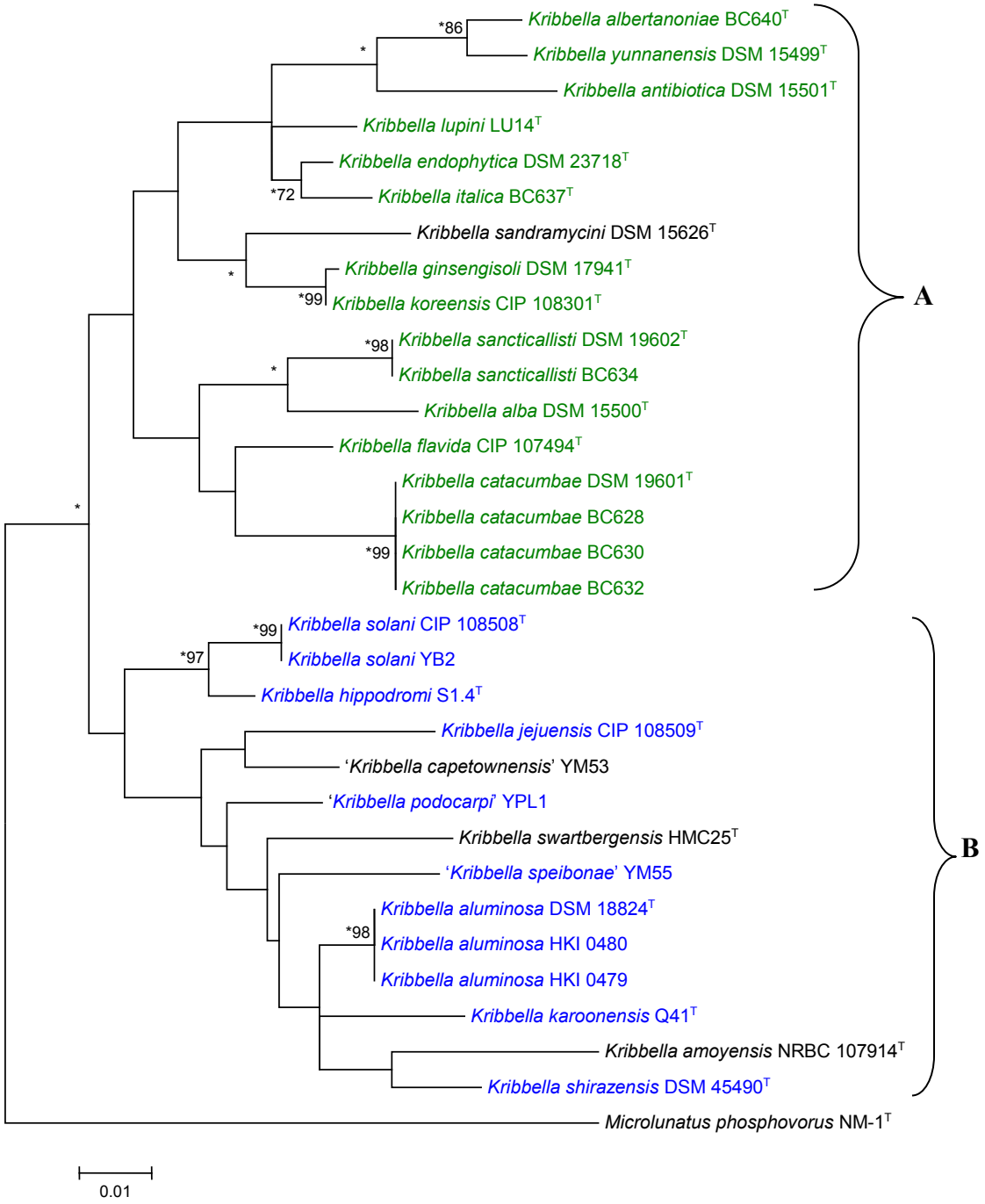
Appendix 4D: *recA* gene phylogenetic tree for the genus *Kribbella.* The tree is based on 519 nt of common sequence and was constructed using the maximum likelihood method. The bootstrap values are based on 1000 resampled datasets and only values greater than 70% are shown. Asterisks indicate the nodes that were also obtained in the maximum-parsimony and neighbour-joining trees. The scale bar represents 2 nt substitutions per 100 nt. *Microlunatus phosphovorus* NM-1T was used as an outgroup.

Appendix 4E (*relA* tree)



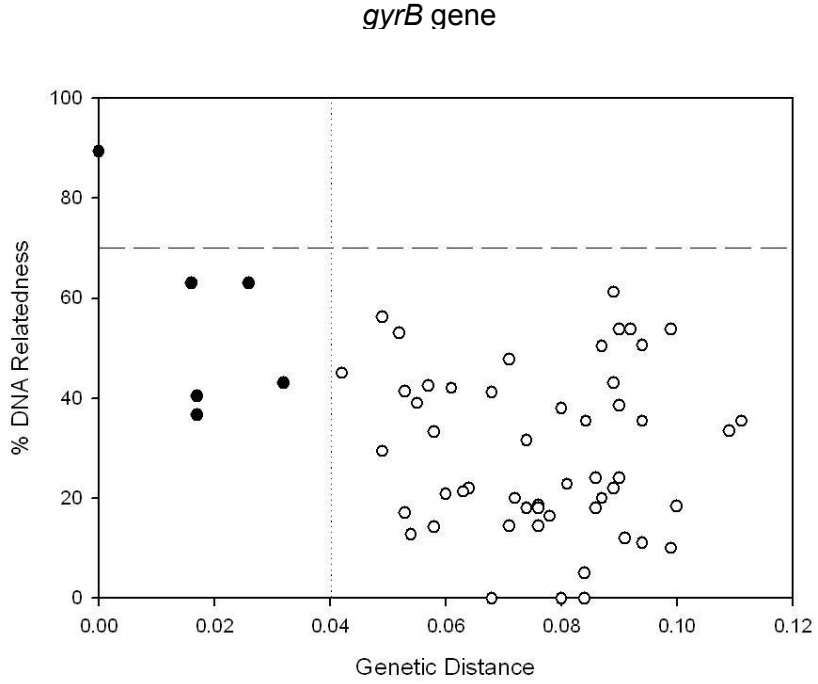
Appendix 4E: *relA* gene phylogenetic tree for the genus *Kribbella.* The tree is based on 1014 nt of common sequence and was constructed using the maximum likelihood method. The bootstrap values are based on 1000 resampled datasets and only values greater than 70% are shown. Asterisks indicate the nodes that were also obtained in the maximum-parsimony and neighbour-joining trees. The scale bar represents 5 nt substitutions per 100 nt. *Microlunatus phosphovorus* NM-1T was used as an outgroup.

Appendix 4F (*atpD* tree)



Appendix 4F: *atpD* gene phylogenetic tree for the genus *Kribbella.* The tree is based on 573 nt of common sequence and was constructed using the maximum likelihood method. The bootstrap values are based on 1000 resampled datasets and only values greater than 70% are shown. Asterisks indicate the nodes that were also obtained in the maximum-parsimony and neighbour-joining trees. The scale bar represents 1 nt substitution per 100 nt. *Microlunatus phosphovorus* NM-1T was used as an outgroup.

Appendix 4G (genetic-distance vs DNA relatedness plot for *gyrB* gene)



Appendix 4G: Graph of *gyrB* gene based genetic distances versus DNA relatedness values for the genus *Kribbella.* The horizontal dashed line indicates the 70% DNA relatedness threshold for distinguishing between genomic species, while the vertical dotted line indicates the proposed genetic distance threshold for distinguishing genomic species in the genus. The points below the 0.04 threshold are pairwise comparisons of *K. hippodromi* S1.4^T vs. *K. solani* CIP 108508^T (40.4% DNA relatedness by DDH) [Everest *et al*., 2011], *K. hippodromi* S1.4^T vs. *K. solani* YB2, (36.6%) [Kirby *et al*., 2010], '*K. podocarpi'* YPL1^T vs *K. karoonensis* Q41^T (43%) [Curtis *et al*., manuscript under revision], '*K. italica*' BC637^T vs *K. lupini* LU14^T and *K. endophytica* DSM 23718^T (both 63%) [Everest *et al*., 2015]. The comparison between *K. solani* CIP 108508^T and *K. solani* YB2 (89.3% DNA relatedness by DDH) [Kirby *et al*., 2010] is the only one under the 0.04 genetic distance threshold and above the 70% DNA relatedness threshold and shows that the strains belong to the same genomic species.

Appendix 4H: Genetic distance table of the genus *Kribbella* using the concatenated *gyrB-rpoB-recA-relA-atpD* genes

	Strains	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
1	*Kribbella podocarpi* YPL1^T																															
2	*Kribbella capetownensis* YM53^T	0.067																														
3	*Kribbella speibonae* YM55^T	0.044	0.059																													
4	*Kribbella alba* DSM 15500^T	0.084	0.084	0.086																												
5	*Kribbella albertanoniae* BC640^T	0.072	0.079	0.076	0.078																											
6	*Kribbella aluminosa* DSM 18824^T	0.044	0.063	0.051	0.086	0.073																										
7	*Kribbella aluminosa* HKI 0479	0.044	0.063	0.051	0.086	0.073	0.000																									
8	*Kribbella aluminosa* HKI 0480	0.044	0.063	0.051	0.086	0.073	0.000	0.000																								
9	*Kribbella amoyensis* NRBC 1007914^T	0.104	0.118	0.109	0.123	0.117	0.111	0.111	0.111																							
10	*Kribbella antibiotica* DSM 15501^T	0.077	0.087	0.079	0.086	0.056	0.075	0.075	0.075	0.128																						
11	*Kribbella catacumbae* DSM 19601^T	0.084	0.084	0.084	0.059	0.066	0.084	0.084	0.084	0.122	0.075																					
12	*Kribbella catacumbae* BC628	0.084	0.084	0.084	0.059	0.066	0.084	0.084	0.084	0.122	0.075	0.000																				
13	*Kribbella catacumbae* BC630	0.084	0.084	0.084	0.059	0.066	0.084	0.084	0.084	0.122	0.075	0.000	0.000																			
14	*Kribbella catacumbae* BC632	0.084	0.084	0.084	0.059	0.066	0.084	0.084	0.084	0.122	0.075	0.000	0.000	0.000																		
15	*Kribbella endophytica* DSM 23718^T	0.075	0.074	0.080	0.065	0.063	0.082	0.082	0.082	0.117	0.071	0.058	0.058	0.058	0.058																	
16	*Kribbella flavida* CIP 107494^T	0.072	0.075	0.072	0.061	0.063	0.076	0.076	0.076	0.106	0.072	0.051	0.051	0.051	0.051	0.046																
17	*Kribbella ginsengisoli* DSM 17941^T	0.078	0.079	0.078	0.060	0.065	0.076	0.076	0.076	0.119	0.073	0.048	0.048	0.048	0.048	0.057	0.052															
18	*Kribbella hippodromi* S1.4^T	0.047	0.068	0.054	0.084	0.076	0.042	0.042	0.042	0.113	0.073	0.082	0.082	0.082	0.082	0.076	0.077	0.076														
19	*Kribbella italica* BC637^T	0.074	0.075	0.077	0.070	0.065	0.081	0.081	0.081	0.114	0.071	0.059	0.059	0.059	0.059	0.025	0.046	0.061	0.074													
20	*Kribbella jejuensis* CIP 108509^T	0.045	0.066	0.054	0.085	0.074	0.042	0.042	0.042	0.113	0.076	0.082	0.082	0.082	0.082	0.078	0.072	0.077	0.046	0.079												
21	*Kribbella karoonensis* Q41^T	0.038	0.070	0.052	0.088	0.080	0.037	0.037	0.037	0.109	0.077	0.088	0.088	0.088	0.088	0.079	0.075	0.084	0.046	0.080	0.040											
22	*Kribbella koreensis* CIP 108301^T	0.073	0.078	0.076	0.063	0.068	0.071	0.071	0.071	0.121	0.074	0.049	0.049	0.049	0.049	0.060	0.054	0.014	0.072	0.064	0.072	0.074										
23	*Kribbella lupini* LU14^T	0.069	0.072	0.075	0.064	0.062	0.078	0.078	0.078	0.111	0.075	0.057	0.057	0.057	0.057	0.025	0.044	0.060	0.069	0.025	0.073	0.076	0.062									
24	*Kribbella sancticallisti* DSM 19602^T	0.072	0.076	0.079	0.052	0.064	0.079	0.079	0.079	0.111	0.070	0.042	0.042	0.042	0.042	0.051	0.046	0.053	0.078	0.053	0.075	0.079	0.055	0.051								
25	*Kribbella sancticallisti* BC634	0.072	0.076	0.079	0.052	0.064	0.079	0.079	0.079	0.111	0.070	0.042	0.042	0.042	0.042	0.051	0.046	0.053	0.078	0.053	0.075	0.079	0.055	0.051	0.000							
26	*Kribbella sandramycini* DSM 15626^T	0.067	0.071	0.069	0.080	0.070	0.070	0.070	0.070	0.121	0.071	0.082	0.082	0.082	0.082	0.070	0.069	0.073	0.069	0.068	0.066	0.074	0.072	0.065	0.077	0.077						
27	*Kribbella solani* CIP 108508^T	0.050	0.067	0.055	0.084	0.077	0.042	0.042	0.042	0.115	0.074	0.082	0.082	0.082	0.082	0.077	0.078	0.076	0.018	0.076	0.048	0.050	0.072	0.072	0.080	0.080	0.069					
28	*Kribbella solani* YB2	0.050	0.067	0.055	0.084	0.077	0.042	0.042	0.042	0.115	0.074	0.082	0.082	0.082	0.082	0.077	0.078	0.076	0.018	0.076	0.048	0.050	0.072	0.072	0.080	0.080	0.069	0.000				
29	*Kribbella shirazensis* DSM 45490^T	0.040	0.062	0.044	0.083	0.072	0.042	0.042	0.042	0.106	0.073	0.085	0.085	0.085	0.085	0.076	0.071	0.076	0.049	0.076	0.045	0.043	0.075	0.072	0.075	0.075	0.065	0.049	0.049			
30	*Kribbella swartbergensis* HMC25^T	0.057	0.053	0.058	0.081	0.078	0.061	0.061	0.061	0.099	0.084	0.079	0.079	0.079	0.079	0.076	0.066	0.075	0.066	0.071	0.061	0.062	0.073	0.072	0.070	0.070	0.074	0.069	0.069	0.058		
31	*Kribbella yunnanensis* DSM 15499^T	0.070	0.077	0.073	0.076	0.033	0.071	0.071	0.071	0.117	0.060	0.064	0.064	0.064	0.064	0.060	0.060	0.057	0.075	0.062	0.073	0.077	0.063	0.061	0.063	0.063	0.062	0.074	0.074	0.070	0.073	
32	*Microlunatus phosphovorus* NM-1^T	0.202	0.204	0.211	0.214	0.209	0.210	0.210	0.210	0.232	0.216	0.215	0.215	0.215	0.215	0.201	0.205	0.206	0.200	0.205	0.204	0.203	0.204	0.201	0.207	0.207	0.205	0.200	0.200	0.210	0.195	0.208

Appendix H. Estimates of Evolutionary Divergence between Sequences

The numbers of base substitutions per site between pairs of sequences are shown. Analyses were conducted using the Kimura 2-parameter model [1]. The analysis involved 32 nucleotide sequences. Codon positions included were 1st+2nd+3rd. All ambiguous positions were removed for each sequence pair before the genetic distance between them was calculated. There were a total of 4099 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 [2].

1. **Kimura M. (1980).** A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* **16**:111-120.
2. **Tamura K., Peterson D., Peterson N., Stecher G., Nei M., and Kumar S. (2011).** MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* **28,** 2731-2739.

Appendix 4I: Genetic distance table for the genus *Kribbella* using only the *gyrB* gene

	Strains	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
1	*Kribbella podocarpi* YPL1^T																															
2	*Kribbella capetownensis* YM53^T	0.081																														
3	*Kribbella speibonae* YM55^T	0.062	0.054																													
4	*Kribbella alba* DSM 15500^T	0.105	0.110	0.120																												
5	*Kribbella albertanoniae* BC640^T	0.095	0.106	0.111	0.098																											
6	*Kribbella aluminosa* DSM 18824^T	0.057	0.085	0.075	0.108	0.098																										
7	*Kribbella aluminosa* HKI 0479	0.057	0.085	0.075	0.108	0.098	0.000																									
8	*Kribbella aluminosa* HKI 0480	0.057	0.085	0.075	0.108	0.098	0.000	0.000																								
9	*Kribbella amoyensis* NRBC 1007914^T	0.073	0.102	0.087	0.094	0.084	0.090	0.090	0.090																							
10	*Kribbella antibiotica* DSM 15501^T	0.080	0.100	0.097	0.111	0.090	0.082	0.082	0.082	0.079																						
11	*Kribbella catacumbae* DSM 19601^T	0.104	0.105	0.110	0.088	0.065	0.108	0.108	0.108	0.087	0.086																					
12	*Kribbella catacumbae* BC628	0.104	0.105	0.110	0.088	0.065	0.108	0.108	0.108	0.087	0.086	0.000																				
13	*Kribbella catacumbae* BC630	0.104	0.105	0.110	0.088	0.065	0.108	0.108	0.108	0.087	0.086	0.000	0.000																			
14	*Kribbella catacumbae* BC632	0.104	0.105	0.110	0.088	0.065	0.108	0.108	0.108	0.087	0.086	0.000	0.000	0.000																		
15	*Kribbella endophytica* DSM 23718^T	0.084	0.088	0.094	0.089	0.076	0.100	0.100	0.100	0.083	0.087	0.080	0.080	0.080	0.080																	
16	*Kribbella flavida* CIP 107494^T	0.072	0.094	0.084	0.084	0.066	0.088	0.088	0.088	0.068	0.074	0.063	0.063	0.063	0.063	0.052																
17	*Kribbella ginsengisoli* DSM 17941^T	0.083	0.093	0.094	0.090	0.077	0.087	0.087	0.087	0.082	0.086	0.063	0.063	0.063	0.063	0.071	0.064															
18	*Kribbella hippodromi* S1.4^T	0.057	0.092	0.075	0.104	0.104	0.057	0.057	0.057	0.083	0.070	0.108	0.108	0.108	0.108	0.091	0.084	0.086														
19	*Kribbella italica* BC637^T	0.078	0.087	0.089	0.094	0.078	0.095	0.095	0.095	0.077	0.083	0.078	0.078	0.078	0.078	0.016	0.051	0.070	0.090													
20	*Kribbella jejuensis* CIP 108509^T	0.041	0.082	0.078	0.101	0.089	0.053	0.053	0.053	0.086	0.079	0.096	0.096	0.096	0.096	0.083	0.071	0.084	0.050	0.084												
21	*Kribbella karoonensis* Q41^T	0.032	0.084	0.065	0.105	0.102	0.045	0.045	0.045	0.080	0.076	0.105	0.105	0.105	0.105	0.086	0.077	0.091	0.049	0.087	0.035											
22	*Kribbella koreensis* CIP 108301^T	0.060	0.088	0.085	0.099	0.086	0.068	0.068	0.068	0.086	0.090	0.074	0.074	0.074	0.074	0.084	0.072	0.042	0.073	0.085	0.060	0.051										
23	*Kribbella lupini* LU14^T	0.082	0.086	0.091	0.089	0.073	0.099	0.099	0.099	0.081	0.095	0.076	0.076	0.076	0.076	0.024	0.054	0.076	0.085	0.026	0.081	0.082	0.085									
24	*Kribbella sancticallisti* DSM 19602^T	0.073	0.094	0.098	0.078	0.068	0.088	0.088	0.088	0.075	0.076	0.049	0.049	0.049	0.049	0.069	0.054	0.067	0.095	0.067	0.079	0.081	0.075	0.065								
25	*Kribbella sancticallisti* BC634	0.073	0.094	0.098	0.078	0.068	0.088	0.088	0.088	0.075	0.076	0.049	0.049	0.049	0.049	0.069	0.054	0.067	0.095	0.067	0.079	0.081	0.075	0.065	0.000							
26	*Kribbella sandramycini* DSM 15626^T	0.064	0.077	0.076	0.100	0.109	0.079	0.079	0.079	0.092	0.094	0.108	0.108	0.108	0.108	0.094	0.089	0.089	0.075	0.089	0.068	0.070	0.084	0.090	0.088	0.088						
27	*Kribbella solani* CIP 108508^T	0.063	0.089	0.082	0.104	0.110	0.058	0.058	0.058	0.085	0.069	0.111	0.111	0.111	0.111	0.099	0.087	0.094	0.017	0.096	0.057	0.057	0.078	0.093	0.100	0.100	0.080					
28	*Kribbella solani* YB2	0.063	0.089	0.082	0.104	0.110	0.058	0.058	0.058	0.085	0.069	0.111	0.111	0.111	0.111	0.099	0.087	0.094	0.017	0.096	0.057	0.057	0.078	0.093	0.100	0.100	0.080	0.000				
29	*Kribbella shirazensis* DSM 45490^T	0.050	0.066	0.060	0.100	0.097	0.061	0.061	0.061	0.074	0.066	0.098	0.098	0.098	0.098	0.072	0.066	0.080	0.060	0.072	0.055	0.049	0.071	0.076	0.080	0.080	0.063	0.060	0.060			
30	*Kribbella swartbergensis* HMC25^T	0.067	0.073	0.062	0.120	0.098	0.076	0.076	0.076	0.071	0.084	0.103	0.103	0.103	0.103	0.092	0.077	0.099	0.081	0.086	0.075	0.071	0.090	0.090	0.090	0.090	0.085	0.085	0.085	0.063		
31	*Kribbella yunnanensis* DSM 15499^T	0.089	0.090	0.103	0.092	0.052	0.089	0.089	0.089	0.083	0.076	0.063	0.063	0.063	0.063	0.065	0.058	0.053	0.094	0.062	0.086	0.098	0.071	0.069	0.060	0.060	0.089	0.099	0.099	0.086	0.094	
32	*Microlunatus phosphovorus* NM-1^T	0.249	0.248	0.253	0.260	0.258	0.259	0.259	0.259	0.249	0.264	0.263	0.263	0.263	0.263	0.253	0.259	0.245	0.238	0.254	0.237	0.241	0.241	0.249	0.262	0.262	0.257	0.242	0.242	0.250	0.229	0.261

Appendixl. Estimates of Evolutionary Divergence between Sequences
The numbers of base substitutions per site between pairs of sequences are shown. Analyses were conducted using the Kimura 2-parameter model [1]. The analysis involved 32 nucleotide sequences. Codon positions included were 1st+2nd+3rd. All ambiguous positions were removed for each sequence pair before the genetic distance between them was calculated. There were a total of 1135 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 [2].

1. **Kimura M.** (**1980**). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* **16**:111-120.

2. **Tamura K., Peterson D., Peterson N., Stecher G., Nei M., and Kumar S. (2011).** MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* **28,** 2731-2739.

Appendix 4J Species description paper for strains YPL1, YM53 and YM55

Kribbella podocarpi sp. nov., isolated from the leaves of a yellowwood tree (*Podocarpus latifolius*) and *Kribbella capetownensis* sp. nov. and *Kribbella speibonae* sp. nov., isolated from soil

Sarah M. Curtis, Iulia Norton, Gareth J. Everest and Paul R. Meyers*

Department of Molecular and Cell Biology, University of Cape Town, Private Bag X3, Rondebosch, 7701, Cape Town, South Africa.

* Corresponding author
Tel: +27 21 650-3261
Fax: +27 21 650-1861
e-mail: paul.meyers@uct.ac.za

Keywords: multilocus sequence analysis, endophyte, family *Nocardioidaceae*

Footnote: The GenBank accession numbers for the 16S rRNA, *atpD*, *gyrB*, *recA*, *relA* and *rpoB* gene sequences of strainYPL1^T are KM382222, KM382207, KM382216, KM382219, KM382210 and KM382213, respectively; those for strain YM53^T are KM382223, KM382208, KM382217, KM382220, KM382211 and KM382214, respectively, and those for strain YM55^{T}are KM382224, KM382209, KM382218, KM382221, KM382212 and KM382215, respectively.

Appendix 4J continued

Abstract

Three actinobacteria strains were isolated from samples collected from the University of Cape Town and the slopes of Devil's Peak, Cape Town, South Africa. Analysis of the 16S rRNA genes showed that the three strains belonged to the genus *Kribbella*. Phylogenetic analyses using the 16S rRNA gene and multilocus sequence analysis using the concatenated gene sequences of the *gyrB*, *rpoB*, *relA*, r*ecA* and *atpD* genes showed that strains YPL1^T and YM55^T were most closely related to the type strains of *Kribbella karoonensis* and *Kribbella shirazensis*, while strain YM53^T was most closely related to the type strains of *Kribbella swartbergensis* and *Kribbella amoyensis.* Based on the *gyrB* and the concatenated *gyrB-rpoB-relA-*r*ecA-atpD* genetic distance analyses, strains YM53^T and YM55^T were shown to be distinct from all *Kribbella* type strains. DNA-DNA hybridisation experiments showed that strain YPL1^T is a distinct genomic species from its closest phylogenetic relative, *K. karoonensis* Q41^T. Physiological comparisons further showed that strains YPL1^T, YM53^T and YM55^T are phenotypically distinct from the type strains of *Kribbella hippodromi, Kribbella solani, Kribbella jejuensis, Kribbella aluminosa, K. karoonensis, K. shirazensis and K. swartbergensis.* Strains YPL1^T, YM53^T and YM55^T are thus presented as the type strains of novel species, for which the names *Kribbella podocarpi* sp. nov.(=DSM XXXXX^T = NRRL B-XXXXX^T), *Kribbella capetownensis* sp. nov.(=DSM XXXXX^T = NRRL B-XXXXX^T) and *Kribbella speibonae* sp. nov.(=DSM XXXXX^T = NRRL B-XXXXX^T), respectively, are proposed.

Appendix 4J continued

Introduction

The genus *Kribbella* was established by Park et al. (1999) and currently consists of 20species with validly published names (Parte 2014). The most recent additions to the genus are *Kribbella endophytica* (Kaewkla and Franco 2013), *Kribbella shirazensis* (Mohammadipanah et al. 2013) and *Kribbella albertanoniae* (Everest et al. 2013). The members of the genus are Gram positive and contain LL-diaminopimelic acid as the characteristic diamino acid in the cell wall peptidoglycan, phosphatidylcholine in the phospholipid profile and MK-9(H_4) as the major menaquinone (Park et al 1999; Sohn et al 2003; Everest et al 2013). Recently, the taxonomy of the genus *Kribbella* was enhanced through phylogenetic analyses based on the *gyrB* gene (Kirby et al. 2010). This work was subsequently expanded to a multilocus sequence analysis (MLSA) of both type and non-type strains in the genus based on the *gyrB*, *rpoB*, *recA*, *relA* and *atpD* genes (Curtis and Meyers 2012). Furthermore, the use of genetic-distance measurements was shown to be a useful way of assessing whether two *Kribbella* strains belong to different genomic species. Both the Kirby et al. (2010) and Curtis and Meyers (2012) studies proposed genetic-distance thresholds that can be used as a guide to determine whether DNA-DNA hybridization (DDH) analysis is required to determine whether a new *Kribbella* isolate represents a unique genomic species. One of the surprising discoveries of these studies was that the type and non-type strains of the same species have identical *gyrB*, *rpoB*, *recA*, *relA* and *atpD* gene sequences (Kirby et al. 2010; Curtis and Meyers 2012), although this analysis was limited to the four species for which non-type strains are available.

While most *Kribbella* type strains have been isolated from soil, two were isolated as endophytes from plants, namely, *Kribbella lupini* (Trujillo et al. 2006) and *K. endophytica* (Kaewkla and Franco 2013). Here we describe the characterisation of three novel members of the genus. Strains YM53^T and YM55^T were isolated from a soil sample obtained from the University of Cape Town, while the endophytic strain, YPL1^T, was isolated from the leaves of the South African National Tree, *Podocarpus latifolius* (yellowwood tree), collected on Devil's Peak, Table Mountain National Park, Cape Town.

Materials and Methods

Isolation and maintenance of the organisms

A soil sample was obtained from a hill near the dam at the University of Cape Town, South Africa. The soil sample was subjected to dry heat pre-treatment at 60°C for 1 hour, after which approximately 0.1 g of soil was added to 1 ml sterile water and vortexed for 1 minute. The soil suspension was then serially diluted and spread-plated onto International *Streptomyces* Project

Appendix 4J continued

medium 2 (ISP 2) (Shirling and Gottlieb 1966). The medium was supplemented with cycloheximide (50 µg/ml) and nalidixic acid (10 µg/ml). The plates were incubated at 30°C for 21 days. The isolates were purified and maintained on ISP2 medium (Shirling and Gottlieb 1966).

The endophytic strain, YPL1^T, was isolated from the leaves of a yellowwood tree (*P. latifolius*) growing on the slopes of Devil's Peak in the Table Mountain National Park (TMNP), Cape Town. The surfaces of the leaves were sterilized in 70% ethanol for 1 minute, after which the leaves were soaked in 1% NaOCl for 3 minutes and rinsed twice with sterile distilled water. The sterilized leaves were prepared for endophyte isolation as described by Kirby and Meyers (2010). The samples were incubated at 30°C for 30 minutes with gentle shaking before the leaf extracts were serially diluted in a ten-fold series in sterile distilled water and spread-plated onto ISP 2, supplemented with cycloheximide and nalidixic acid, as for the soil sample. The plates were incubated at 30°C for 21 days. The isolate was purified and maintained on ISP2 medium (Shirling and Gottlieb 1966).

DNA extraction, PCR amplification and sequencing

Genomic DNA was extracted as described by Everest et al. (2011). The 16S rRNA genes were amplified as described by Cook and Meyers (2003), the *gyrB* genes as described by Kirby et al. (2010) and the *atpD*, *recA*, *relA* and *rpoB* genes as described by Curtis and Meyers (2012). Approximately 500ng of template DNA was used in the PCR amplification of the 16S rRNA and *gyrB* genes, with 1000ng of DNA being used for the amplification of the *atpD*, *recA*, *relA* and *rpoB* genes. The PCR products were purified using an MSB® Spin PCRapace kit (STRATEC Molecular, Berlin, Germany) and sequenced by Macrogen Inc., Seoul, South Korea. Sequence assembly was performed using DNAMAN version 5.2.9 (LynnonBioSoft). For phylogenetic analyses, sequences were aligned using Muscle (Edgar 2004) and phylogenetic trees were constructed using the maximum-likelihood, maximum-parsimony and neighbour-joining methods with 1000 bootstrap replicates in MEGA version 5.05 (Tamura et al. 2011). Kimura's 2-parameter model was used to calculate the *gyrB* and concatenated-gene genetic distances (Kimura 1980).

The gene sequences determined for the three strains in this study were deposited in the GenBank database. The accession numbers for the 16S rRNA, *atpD*, *gyrB*, *recA*, *relA* and *rpoB* gene sequences of strainYPL1^T are KM382222, KM382207, KM382216, KM382219, KM382210 and KM382213, respectively; those for strain YM53^T are KM382223, KM382208, KM382217, KM382220, KM382211 and KM382214, respectively, and those for strain YM55^{T}are KM382224, KM382209, KM382218, KM382221, KM382212 and KM382215, respectively.

Appendix 4J continued

Morphological and physiological characteristics

All ISP media were prepared according to the methods of Shirling and Gottlieb (1966). Morphological properties were determined on ISP 2 and inorganic salts-starch agar (ISP 4). Physiological tests were carried out as described by Williams et al. (1989). All plates were incubated at 30°C, for the recommended incubation periods, unless otherwise stated. Carbon source utilization was tested as per the methods of Shirling and Gottlieb (1966). All carbon sources were filter-sterilized and tested at a final concentration of 1% (w/v), with the exception of sodium acetate and sodium citrate, which were tested at 0.1% (w/v). Nitrogen source utilization was performed as per Williams et al. (1989). All nitrogen sources were filter sterilized and tested at a final concentration of 0.1% (w/v). NaCl tolerance was determined on ISP 2 medium incubated for 14 days. Growth at pH 4.3, 7 and 10 was determined on Bennett's medium (Atlas 2004) incubated for 14 days.

Chemotaxonomic characterization

Chemotaxonomic analysis of the whole cell sugar pattern and the isomer of diaminopimelic acid in the peptidoglycan was conducted as described by Hasegawa et al. (1983). The solvent system used for the whole-cell sugar analysis was ethyl acetate-pyridine-distilled water (100:35:25 v/v). The phospholipid pattern was determined as described by Komagata and Suzuki (1987) and Minnikin et al. (1984) using α-naphthol, ninhydrin and molybdenum blue reagents. Analysis of respiratory quinones was carried out by the identification service at the DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen), Braunschweig, Germany.All the chemotaxonomic analyses were performed on freeze-dried cells of cultures grown in ISP 2 broth at 30°C for 3 days with shaking. Fatty acid analysis was performed as a service by the BCCM/LMG Culture Collection, as per the recommendations of the commercial identification system MIDI (Microbial Identification System, Inc., Delaware, U.S.A.; MIDI Sherlock version 3.10; database TSBA 50 rev 5.0), on cells grown at 28°C for 3 days on Tryptic Soy Agar (BBL 11768).

Results and Discussion

The closest phylogenetic relatives of strain YPL1^T, based on 16S rRNA gene sequence similarities (as calculated by EzTaxon-e; Kim et al.2012) were, *Kribbella karoonensis* DSM 17344^T (99.8%), *Kribbella swartbergensis* HMC25^T (99.6%), *K. shirazensis* UTMC 693^T (99.5%) and *Kribbella aluminosa* DSM 18824^T (99.1%), while those for strain YM55^T were *K. karoonensis* DSM 17344^T (99.3%), *K. shirazensis* UTMC 693^T (99.2%), *Kribbella hippodromi* S1.4^T (99.1%) and *K. aluminosa* HKI 0478^T (99.0%). For strain YM53^T, the highest 16S rRNA gene sequence

Appendix 4J continued

similarities were to *K. albertanoniae* BC640^T (99.0%), *K. hippodromi* S1.4^T (99.0%), *Kribbella sandramycini* ATCC 39419^T (99.0%), *K. swartbergensis* DSM 17345^T (98.9%), *Kribbella sancticallisti* BC633^T (98.8%), *K. karoonensis* Q41^T (98.8%) and *Kribbella flavida* DSM 17836^T (98.7%). The 16S rRNA gene sequence similarities between the three strains (as calculated by DNAMAN) were as follows: YPL1^T and YM53^T: 98.4%; YPL1^T and YM55^T: 99.4%; and YM55^T and YM53^T: 98.6%. In the 16S rRNA gene maximum-likelihood phylogenetic tree (Figure 1), strains YPL1^T andYM55^T grouped with the type strains of *K. karoonensis*, *K. shirazensis*, *K. aluminosa* and *Kribbella jejuensis* with strong bootstrap support (90%). The closest relatives of strain YPL1^T were *K. karoonensis* Q41^T and *K. shirazensis* UTMC 693^T.

MLSA with the concatenated *gyrB*, *rpoB*, *recA*, *relA* and *atpD* genes (Figure 2) showed that strains YPL1^T and YM55^T formed a well-supported cluster (bootstrap value 90%) with the type strains of *K. shirazensis*, *K. karoonensis*, *K. aluminosa*, *K. jejuensis*, *K. hippodromi* and *Kribbella solani* (cluster B of Curtis and Meyers 2012). Strain YPL1^T was loosely associated with *K. karoonensis* Q41^T in cluster B (bootstrap value 57%). Strain YM53^T formed a deep branch in the MLSA tree with low bootstrap support of 40% (Figure 2).

Genetic distance values, based on a 1135-nt fragment of the *gyrB* gene, between strains YPL1^T, YM53^T and YM55^T and all *Kribbella* type strains with validly-published names were calculated to assess the likelihood of the three strains representing novel genomic species. For strain YM55^T, the *gyrB*-based genetic distance values ranged from 0.054 to 0.120, while those for strain YM53^T ranged from 0.054 to 0.110. The *gyrB* genetic distance between the three strains were as follows: YPL1^T and YM53^T: 0.081; YPL1^T and YM55^T: 0.062; and YM55^T and YM53^T: 0.054. These values are all above the 0.04 *gyrB* genetic distance threshold proposed to distinguish between genomic species in the genus (Kirby et al 2010). The *gyrB* based genetic distance values for strain YPL1^T ranged from 0.032 to 0.105. The only value under the 0.04 threshold (0.032) was between strain YPL1^T and *K. karoonensis* Q41^T.

The concatenated *gyrB-rpoB-recA-relA-atpD* gene sequences (4099 nt) were then used to calculate genetic distances (Curtis and Meyers 2012). The values for strain YM53^T ranged from 0.053-0.118, which are all above the threshold of 0.04 for distinguishing between genomic species. The values for strain YM55^T ranged from 0.044-0.109 when compared to *Kribbella* type strains. Strain YPL1^T had genetic distance values ranging from 0.038-0.104, the lowest value (0.038) corresponding to the comparison of strain YPL1^T with *K. karoonensis* Q41^T. The genetic distance between strain YPL1^T and strain YM55^T, which were in the same cluster (Figure 2) was 0.044, above the 0.04 threshold. Due to the genetic distance values for strain YPL1^T being under

Appendix 4J continued

the threshold for both the *gyrB* gene and concatenated-gene genetic distances, further analysis was performed in order to determine if it belongs to a distinct species. Genetic distance values are shown in Supplementary Tables S2 (*gyrB*) and S3 (*gyrB-rpoB-recA-relA-atpD* concatenated genes).

DDH analysis between strain YPL1^T and *K. karoonensis* Q41^T showed a DNA-relatedness of 43% ± 13%. This result is under the 70% DNA relatedness threshold for genomic species proposed by Wayne et al (1987) and thus strain YPL1^T is a distinct species from *K. karoonensis.* DDH experiments were not performed on strains YM55^T and YM53^T, as the genetic distance values based on the *gyrB* gene and concatenated *gyrB-rpoB-recA-relA-atpD* gene were above the thresholds for distinguishing genomic species in the genus *Kribbella* (Kirby et al 2010; Curtis and Meyers 2012).

The results of the phenotypic characterisation of all three strains are presented in Table 1 and in the species descriptions. All three strains were phenotypically distinct from their closest phylogenetic relatives. With regards to differences between the closely related strains YM55^T and YPL1^T, strain YM55^T was able to utilize inulin as a sole carbon source and L-phenylalanine as a sole nitrogen source, whereas strain YPL1^T could not use these compounds. Strain YPL1^T was able to weakly utilize sodium acetate as a sole carbon source, whereas strain YM55^T could not utilize it. Strain YM55^T was also able to grow weakly in the presence of 4% (w/v) NaCl.

Chemotaxonomic analysis of strains YPL1^T, YM53^T and YM55^T showed that they have chemotaxonomic characteristics consistent with membership of the genus *Kribbella*. All three strains contained LL-DAP as the diagnostic diamino acid in the peptidoglycan, had a phospholipid pattern characterised by the presence of phosphatidylcholine (and the absence of phosphatidylethanolamine) and contained MK-9(H_4) as the major menaquinone and anteiso-$C_{15:0}$ as the predominant fatty acid. Detailed fatty-acid profiles for the three strains are shown in Supplementary Table S1.

Description of *Kribbella podocarpi* sp. nov.

Kribbella podocarpi [po.do.car'pi. N.L. gen. n. *podocarpi*, of *Podocarpus*, isolated from a yellowwood tree (*Podocarpus latifolius*)]

Cells are Gram positive, oxidase and catalase positive. Colonies are beige and wrinkly with irregular edges. Grows at 30°C and 37°C and at pH 4.3, pH 7 and pH 10. Does not grow at 45°C.Grows in the presence of up to 3% (w/v) NaCl. No diffusible pigments are produced.

Appendix 4J continued

Melanin is not produced on peptone-yeast extract-iron agar (ISP 6) or tyrosine agar (ISP 7). Nitrate is reduced to nitrite and H_2S is produced. D(+)–glucose, L(+)-arabinose, D(-)-fructose, *myo*-inositol, D(-)-mannitol, D(+)-mannose, raffinose, L(+)-rhamnose, sucrose, D(+)-xylose, melibiose, α-lactose, salicin, adonitol and cellobiose are utilized as sole carbon sources, while sodium acetate is weakly utilized. Unable to use inulin, and sodium citrate as sole carbon sources. Utilizes L-asparagine, L-arginine, L-histidine, potassium nitrate, L-threonine, DL-α-amino-n-butyric acid, L-cysteine, L-4-hydroxyproline, L-serine and L-valine as sole nitrogen sources. Utilizes L-methionine weakly but is unable to utilize L--phenylalanine as a sole nitrogen source. Starch is weakly hydrolysed and pectin is not hydrolysed. Adenine, casein, gelatin, hypoxanthine and Tween 80 are degraded. L-Tyrosine and urea are weakly degraded. Guanine, xanthine and xylan are not degraded. The cell-wall peptidoglycan contains LL-diaminopimelic acid and glycine. Major fatty acids (>10%) are iso-$C_{15:0}$ and anteiso-$C_{15:0}$. For the type strain, the major menaquinone is MK-9(H_4) (90%) with minor quantities of MK-9 (2%), MK-9(H_2) (4%) and MK-9(H_6) (2%). The polar lipid profile includes diphosphatidylglycerol (DPG), phosphatidylglycerol (PG), phosphatidylcholine (PC) and phosphatidylinositol (PI) with an additional six unidentified phospholipids and two unidentified aminolipids.

The type strain, YPL1^T (=DSM XXXXX^T = NRRL B-XXXXXX^T), was isolated from the leaves of a yellowwood tree, *Podocarpus latifolius*, growing on the slopes of Devil's Peak, Table Mountain National Park, Cape Town, South Africa.

Description of *Kribbella capetownensis* sp. nov.

Kribbella capetownensis [cape.town'en.sis. N.L. fem. adj. *capetownensis*, of or belonging to Cape Town]

Cells are Gram positive, oxidase and catalase positive. Colonies are beige and wrinkly with irregular edges. Grows at 30°C (weakly at 37°C) and at pH 7 and at pH 10. Grows in the presence of up to 3% (w/v) NaCl. Does not grow at pH 4.3 or at 45°C. No diffusible pigments are produced. Melanin is not produced on peptone-yeast extract-iron agar (ISP 6) or tyrosine agar (ISP 7). Nitrate is weakly reduced to nitrite and H_2S is produced. D(+)-glucose, D(-)-fructose, *myo*-inositol, D(-)-mannitol, D(+)-mannose, raffinose, L(+)-rhamnose, sucrose, D(+)-xylose, melibiose, α-lactose, adonitol and cellobiose are utilized as sole carbon sources, with weak growth on L(+)-arabinose and salicin. Unable to use inulin, sodium acetate and sodium citrate as sole carbon sources. Utilizes L-asparagine, L-arginine, potassium nitrate, DL-α-amino-n-butyric acid, L-cysteine and L-4-hydroxyproline as sole nitrogen sources, with weak growth on L-methionine, L-threonine, L-serine and L-phenylalanine. Unable to utilize L-histidine and L-valine

Appendix 4J continued

as sole nitrogen sources. Starch is hydrolysed and pectin is not hydrolysed. Adenine, casein, gelatin, hypoxanthine, xanthine, L-tyrosine and Tween 80 are degraded. Guanine, urea and xylan are not degraded. Whole cell hydrolysates contain LL-diaminopimelic acid. Major fatty acids (>10%) are iso-$C_{15:0}$, anteiso-$C_{15:0}$ and iso-$C_{16:0}$. For the type strain, the major menaquinone is MK-9(H_4) (77%) with minor quantities of MK-8 (8%), MK-7(H_4) (7%) and MK-9(H_6) (7%). The polar lipid profile includes diphosphatidylglycerol (DPG), phosphatidylglycerol (PG), phosphatidylcholine (PC), phosphatidylinositol (PI), and four unidentified phospholipids and one unidentified aminolipid.

The type strain, YM53^T (=DSM XXXXX^T = NRRL B-XXXXXX^T), was isolated from soil obtained from the University of Cape Town.

Description of *Kribbella speibonae* sp. nov.

Kribbella speibonae [spe.i.bo'na.e. L. n. *spes -ei*, hope; L. adj. *bonus*, good; N.L. gen. n. *speibonae*, of good hope, to indicate Cape Town, the Cape of Good Hope, South Africa, the geographical location from which the type strain was isolated]

Cells are Gram positive, oxidase and catalase positive. Colonies are beige and wrinkly with irregular edges. Grows at 30°C and 37°C and at pH 4.3, pH 7 and pH 10. Grows weakly in the presence of 4% (w/v) NaCl, but not at 7% (w/v) NaCl or at 45°C. No diffusible pigments are produced. Melanin is not produced on peptone-yeast extract-iron agar (ISP 6) or tyrosine agar (ISP 7). Nitrate is reduced to nitrite and H_2S is produced. D(+)-glucose, D(-)-fructose, *myo*-inositol, D(-)-mannitol, D(+)-mannose, raffinose, L(+)-rhamnose, sucrose, D(+)-xylose, melibiose, α-lactose, adonitol, L(+)-arabinose, inulin, salicin and cellobiose are utilized as sole carbon sources. Unable to use sodium acetate and sodium citrate as sole carbon sources. Utilizes L-asparagine, L-arginine, L-histidine, potassium nitrate, L-threonine, DL-α-amino-n-butyric acid, L-4-hydroxyproline, L-serine, L-valine and L-phenylalanine as sole nitrogen sources, with weak growth on L-methionine and L-cysteine. Starch is hydrolysed and pectin is not hydrolysed. Adenine, casein, gelatin, hypoxanthine, urea, L-tyrosine and Tween 80 are degraded. Guanine, xanthine and xylan are not degraded. Whole cell hydrolysates contain LL-diaminopimelic acid. Major fatty acids (>10%) are iso-$C_{15:0}$ and anteiso-$C_{15:0}$. For the type strain, the major menaquinone is MK-9(H_4) (85%) with minor quantities of MK-9(H_6) (11%), and MK-9(H_8) (4%). The polar lipid profile includes diphosphatidylglycerol (DPG), phosphatidylglycerol (PG), phosphatidylcholine (PC), phosphatidylinositol (PI), six unidentified phospholipids and three unidentified aminolipids.

The type strain, YM55^T (=DSM XXXXX^T = NRRL B-XXXXXX^T), was isolated from soil obtained from the University of Cape Town.

Appendix 4J continued

Acknowledgements

S.M.C. holds a National Research Foundation (NRF) Innovation Doctoral Scholarship. Opinions expressed and conclusions arrived at are those of the authors and are not necessarily to be attributed to the NRF. G.J.E. held a Claude Leon Foundation Postdoctoral Fellowship. This work was supported by research grants to P.R.M. from the National Research Foundation (grant numbers 81014 and 85476) and the University Research Committee (University of Cape Town).

Conflict of Interest

The authors declare that they have no conflict of interest.

Appendix 4J continued

References

Atlas RM (2004)*Handbook of Microbiological Media (3rd edition)*. Boca Raton, Fl: CRC Press.

Cook A E and Meyers P R (2003) Rapid identification of filamentous actinomycetes to the genus level using genus-specific 16S-rRNA gene restriction fragment patterns. Int J Syst Evol Microbiol.53: 1907-1915.

Curtis S M and Meyers P R (2012) Multilocus sequence analysis of the actinobacterial genus *Kribbella*. Syst Appl Microbiol. 35: 441-446.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32: 1792-1797.

Everest GJ, Cook AE, Kirby BM, Meyers PR (2011) Evaluation of the use of *recN* sequence analysis in the phylogeny of the genus *Amycolatopsis*. Antonie van Leeuwenhoek. 100: 483-496.

Everest GJ, Curtis SM, De Leo F, Urzì C, and Meyers PR (2013) *Kribbella albertanoniae sp. nov*., isolated from a Roman catacomb, and emended description of the genus *Kribbella*. Int J Syst Evol Microbiol. 63: 3591-3596.

Hasegawa T, Takizawa M and Tanida S (1983) A rapid analysis for chemical grouping of aerobic actinomycetes. J Gen ApplMicrobiol.29: 319-322.

Kaewkla O and Franco C M M (2013) *Kribbella endophytica* sp. nov., an endophytic actinobacterium isolated from the surface-sterilized leaf of a native apricot tree. Int J Syst Evol Microbiol.63: 1249 – 12530.

Kim OS, Cho YJ, Lee K, Yoon SH, Kim M, Na H, Park SC, Jeon YS, Lee JH, Yi H, Won S, Chun J (2012) Introducing EzTaxon-e: a prokaryotic 16S rRNA Gene sequence database with phylotypes that represent uncultured species. Int J Syst Evol Microbiol. 62: 716–721.

Kimura M (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol.16: 111-120.

Kirby BM, Everest GJ and Meyers PR (2010) Phylogenetic analysis of the genus *Kribbella* based on the *gyrB* gene: proposal of a *gyrB*-sequence threshold for species delineation in the genus *Kribbella*. Antonie van Leeuwenhoek. 97: 131-142.

Kirby BM and Meyers PR (2010) *Micromonospora tulbaghiae* sp. nov., isolated from the leaves of wild garlic, *Tulbaghia violacea*. Int J Syst Evol Microbiol.60: 1328-1333.

Appendix 4J continued

Komagata K. and Suzuki KL (1987) Lipid and cell-wall analysis in bacterial systematics.Methods Microbiol. 19: 161-207.

Minnikin DE, O'Donnell AG, Goodfellow M, Alderson G, Athalye M, Schaal A. and Parlett JH (1984) An integrated procedure for the extraction of bacterial isoprenoid quinines and polar lipids. J Microbiol Methods. 2: 233-241.

Mohammadipanah F, Hamedi J, Göker M, Fiebig A, Pukall R, Spröer C and Klenk HP (2013) *Kribbella shirazensis*sp.nov., isolated from soil of Shiraz in Iran. Int J Syst Evol Microbiol. 63: 3369-3374.

Park YH, Yoon JH, Shin YK, Suzuki K, Kudo T, Seino A, Kim H, Lee JS and Lee ST (1999) Classification of '*Nocardioides fulvus*' IFO 14399 and *Nocardioides* sp. ATCC 39419 in *Kribbella* gen. nov., as *Kribbella flavida* sp. nov.and *Kribbella sandramycini* sp. nov.Int J Syst Bacteriol.49: 743-752.

Parte AC (2014) List of prokaryotic names with standing in nomenclature.(http://www.bacterio.net/kribbella.html) Accessed August 2014.

Shirling EB and Gottlieb D (1966) Methods for characterization of *Streptomyces* species.Int J Syst Bacteriol.16: 313-340.

Sohn K, Hong SG, Bae KS and Chun J (2003) Transfer of *Hongia koreensis* Lee *et al.* 2000 to the genus *Kribbella* Park *et al*. 1999 as *Kribbella koreensis* comb. nov.Int J Syst Evol Microbiol. 53: 1005–1007.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M and Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods Mol Biol Evol. 28: 2731-2739.

Trujillo, M. E., Kroppenstedt, R. M., Schumann, P., & Martinez-Molina, E. (2006).*Kribbella lupini* sp. nov., isolated from the roots of *Lupinus angustifolius*. Int J Syst Evol Microbiol 56: 407-411.

Wayne LG, Brenner DJ, Colwell RR, Grimont P. AD, Kandler O, Krichevsky MI, Moore LH, Moore WEC, Murray RGE, Stackebrandt E, Starr MP and Trüper HG (1987) Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. Int J Syst Bacteriol. 37: 463-464.

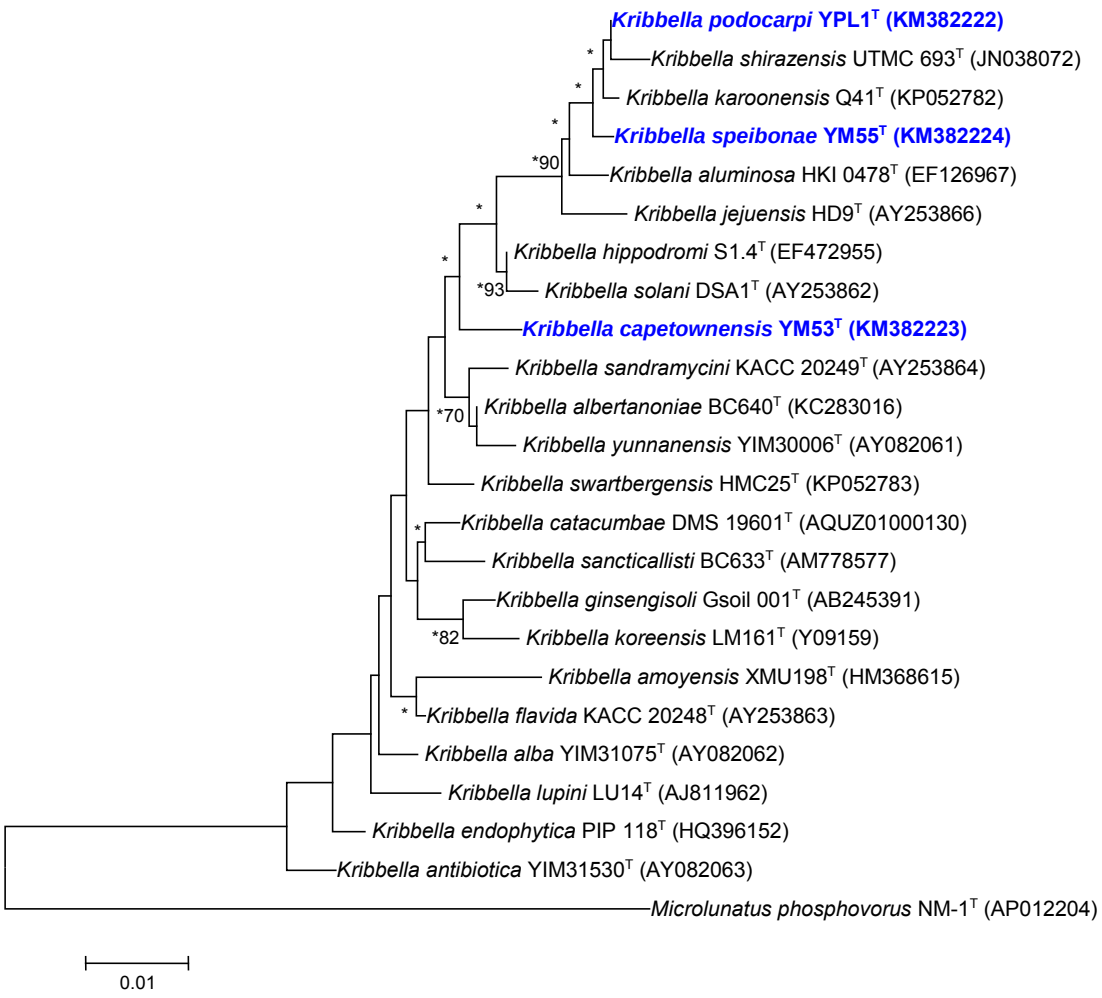
Williams ST, Goodfellow M and Alderson G (1989) Genus *Streptomyces* Waksman and Henrici 1943, 339[AL].In Bergey's Manual of Systematic Bacteriology, vol. 4, pp. 2452-2492.Edited by ST Williams, ME Sharpe and JG Holt. Baltimore: Williams & Wilkins.

Appendix 4J continued

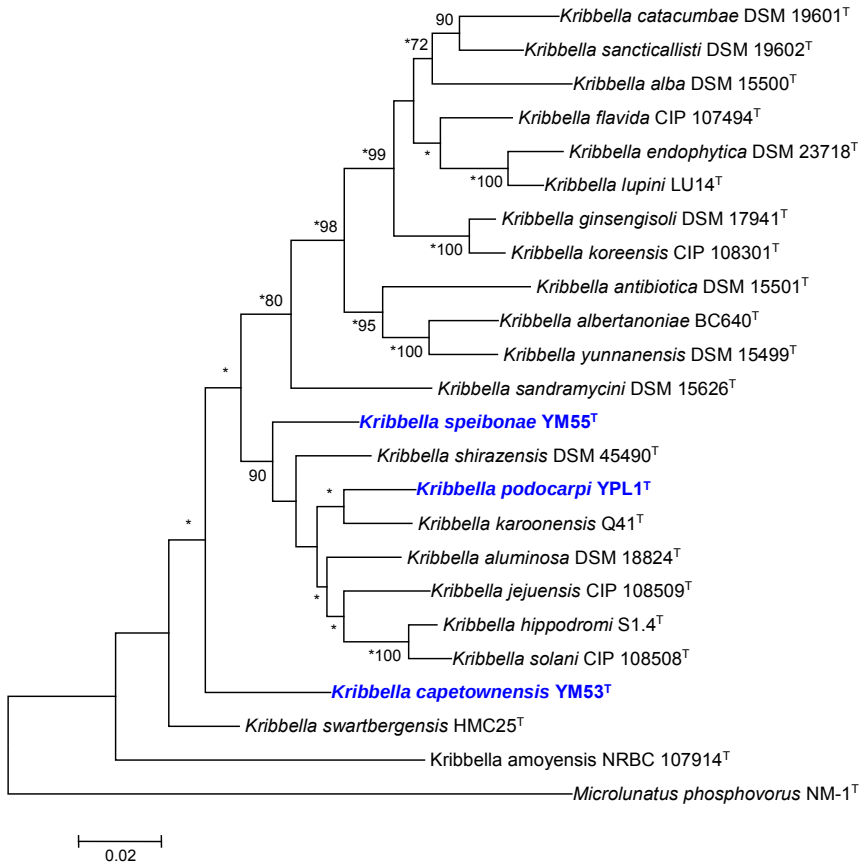
Figure 1.16S rRNA gene phylogenetic tree for the genus *Kribbella*.The tree is based on 1347 nt of common sequence and was constructed using the maximum-likelihood method. The bootstrap values are based on 1000 resampled datasets and only values ≥70% are shown. Asterisks indicate the nodes that were also obtained in the maximum-parsimony and neighbour-joining trees. The scale bar represents 1 nt substitution per 100nt. *Microlunatus phosphovorus* NM-$1^T$was used as an outgroup.

Figure 2. The concatenated *gyrB-rpoB-recA-relA-atpD* gene phylogenetic tree for the genus *Kribbella*. The tree is based on 4099 nt of common sequence and was constructed using the maximum-likelihood method. The bootstrap values are based on 1000 resampled datasets and only values ≥70% are shown. Asterisks indicate the nodes that were also obtained in the maximum-parsimony and neighbour-joining trees. The scale bar represents 2 nt substitutions per 100nt. *Microlunatus phosphovorus* NM-1^T was used as an outgroup.

Appendix 4J continued
Figure1. 16S rRNA gene tree



Appendix 4J continued
Figure 2. *gyrB-rpoB-recA-relA-atpD*



Appendix 4J continued

Table 1. Differential phenotypic characteristics between strains YPL1^T, YM53^T and YM55^T and closely related *Kribbella* type strains. All data were determined in this study. Symbols: ++ strongly positive; +, positive; w+, weak positive; -, negative.
Strain 1, *K. aluminosa* DSM 18824^T; 2, *K. hippodromi* S1.4^T; 3, *K. jejuensis* CIP 108509^T; 4, *K. karoonensis* Q41^T; 5, *K. solani* CIP 108508^T; 6, *K. swartbergensis* HMC25^T; 7, *K. flavida* CIP 107494^T; 8, *K. sancticallisti* DSM 19602^T; 9, *K. shirazensis* DSM 45490^T; 10, *K. albertanoniae* BC640^T.

Test	**YPL1^T**	**YM53^T**	**YM55^T**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
H_2S	+	+	+	+	+	+	+	+	+	-	+	+	+
Growth at:													
4% NaCl	-	-	w+	+	+	w+	+	+	+	w+	-	-	-
pH 4.3	+	-	+	+	+	+	+	+	-	-	-	-	w+
pH 10	+	+	+	+	+	+	+	+	+	++	+	-	w+
37°C	+	w+	+	w+	+	+	+	-	+	+	+	+	w+
45°C	-	-	-	-	-	-	-	-	+	-	-	w+	-
Degradation of:													
Adenine	+	+	+	+	+	+	+	+	+	-	+	+	+
Guanine	-	-	-	-	-	-	-	-	-	-	-	-	w+
Hypoxanthine	+	+	+	+	+	+	+	+	+	-	+	+	+
Xanthine	-	+	-	-	-	-	-	-	-	-	-	-	-
L-tyrosine	w+	+	+	+	+	+	-	+	w+	+	+	-	-
Xylan	-	-	-	-	-	-	-	-	-	-	-	-	+
Urea	w+	-	+	+	-	+	+	-	-	-	-	+	-
Utilization as sole Nitrogen source													
L-histidine	++	-	++	++	++	w+	++	++	+	++	++	++	+
DL-α-Amino- *n*-butyric acid	+	+	++	++	++	++	++	++	+	+	-	++	w+
L-valine	++	-	++	++	++	++	+	++	+	+	+	++	++
L-methionine	w+	w+	w+	-	+	+	-	++	-	+	-	+	w+
L-phenylalanine	-	w+	+	-	w+	+	-	++	-	-	++	+	+
Utilization as sole Carbon source													
Inulin	-	-	+	+	+	+	-	-	-	-	-	-	-
L(+)-Rhamnose	+	+	+	+	+	+	+	+	+	w+	-	-	+
Sodium acetate	w+	-	-	-	w+	w+	- w+	w+	-	-	-	-	w+
Sodium citrate	-	-	-	-	w+	w+	-	w+	-	-	-	-	w+

Appendix 4K species description paper for strain BC637

Description of *Kribbella italica* sp. nov., isolated from a Roman catacomb

Gareth J. Everest[1], Sarah M. Curtis[1], Filomena De Leo[2], Clara Urzì[2] and Paul R. Meyers[1]*

[1]Department of Molecular and Cell Biology, University of Cape Town, Private Bag X3, Rondebosch, 7701, Cape Town, South Africa.
[2]Department of Biological and Environmental Sciences, University of Messina, Viale F. Stagnod'Alcontres 31, I-98166 Messina, Italy.

Running Title: *Kribbella italica* sp. nov.

Keywords: nocardioform actinobacteria, multilocus sequence analysis, biodeteriogenic biofilm

* Corresponding author
Tel: +27 21 650-3261
Fax: +27 21 650-1861
e-mail: paul.meyers@uct.ac.za

Footnote: The GenBank accession numbers for the 16S rRNA, *atpD*, *gyrB*, *recA*, *relA* and *rpoB* gene sequences of strainBC637^T are KJ875927,KJ875925, KJ875917, KJ875921, KJ875923 and KJ875919, respectively.

Appendix 4K continued

Summary:

A novel actinobacterium, strain BC637^T, was isolated from a biodeteriogenic biofilm sample collected in 2009 in the Saint Callixstus Roman catacombs. The strain was found to belong to the genus *Kribbella* by analysis of the 16S rRNA gene. Phylogenetic analysis using the 16S rRNA gene and the *gyrB*, *rpoB*, *relA*, *recA* and *atpD* concatenated gene sequences showed that strain BC637^T was most closely related to the type strains of *Kribbella lupini* and *Kribbella endophytica*. DNA-DNA hybridisation experiments confirmed that strain BC637^T is a distinct genomic species from its closest phylogenetic relatives *K. endophytica* DSM 23718^T (63% DNA relatedness) and *K. lupini* LU 14^T (63% DNA relatedness). Physiological comparisons showed that strain BC637^{T}is phenotypically distinct from the type strains of *K. endophytica* and *K. lupini*. Strain BC637^T is thus presented as the type strain of a novel species, for which the name *Kribbella italica* sp.nov. is proposed (= DSM 28967^T = NRRL B-59155^T).

Appendix 4K continued

Proposed in 1999 by Park *et al.*, the genus *Kribbella* contains nocardioform actinobacteria with LL-diaminopimelic acid in the cell wall peptidoglycan. The genus description was recently emended by Everest *et al.* 2013. Another recent development in the genus was the publication of a multilocus sequence analysis (MLSA) method, which can be used to assess whether an isolate belongs to a new species without the need for DNA-DNA hybridisation (DDH) experiments (Curtis & Meyers, 2012). The MLSA study extended previous phylogenetic analysis of *Kribbella* strains using the *gyrB* gene alone (Kirby *et al*., 2010).With the description of *Kribbella albertanoniae* (Everest *et al.*, 2013), *Kribbella endophytica* (Kaewkla& Franco, 2013) and *Kribbella shirazensis* (Mohammadipanah *et al.* 2013), there are twenty *Kribbella* species with validly-published names (Parte, 2014). Here we describe the characterisation of a novel member of the genus, isolated from the same site as the previously described species *K. albertanoniae*, *Kribbella catacumbae* and *Kribbella sancticallisti* in the Saint Callixstus catacomb in Rome (Everest *et al.*, 2013; Urzì *et al.*, 2008).

Ocean cubicle (CSC13) is located inside the Saint Callixstus catacomb in Rome (Italy) and was extensively studied under the European Community financed project CATS (Cyanobacteria ATtackRockS). During this project, an interdisciplinary group of researchers investigated the damage caused by the growth of microbial communities, mainly cyanobacteria-containing biofilms, as well as white-greyish patinas, as described by Albertano *et al*. (2003). At the end of the project, the Ocean cubicle was closed to the public and blue light was used for illumination to control the growth of photosynthetic microorganisms. Strain BC637^T was isolated in 2009 from a white biofilm on a frescoed surface during a periodic monitoring campaign.

Site CSC13 harboured different species of *Kribbella*, among which were *K. sancticallisti* strain BC633^T, *K. catacumbae* strain BC631^T (both isolated before the microclimatic changes induced by the blue light), and *K. albertanoniae* strain BC640^T isolated in 2009, but from a different

Appendix 4K continued

sampling site than strain BC637^T. Figure S1 shows the sampling locations in the CSC13 cubicle from which the new *Kribbella* species were isolated *(*Urzì *et al*., 2008; Everest *et al*., 2013; the present study). All four of the new species were associated with white/greyish biofilms, except strain BC637^T, which was isolated from a green patina.

Sampling was carried out by using the non-destructive adhesive tape technique (Fungi Tape™, DID, Milan, Italy; Urzì & De Leo, 2001). Growth of colonies (quantified as the number of CFU per cm^2 of adhesive tape) was carried out on R2A medium (Reasoner & Geldreich, 1985) at 28°C for 15 days. Ten to 20 colonies were randomly selected, preliminarily characterized after transferring to TSA (Tryptic Soy Agar, BBL, USA) and subsequently maintained on yeast extract-malt extract agar (International *Streptomyces* Project (ISP) medium 2) (Shirling & Gottlieb, 1966).

Genomic DNA was extracted as described by Everest & Meyers (2008). The 16S rRNA gene was amplified as described by Cook & Meyers (2003), the *gyrB* gene as described by Kirby *et al.* (2010) and the *atpD*, *recA*, *relA* and *rpoB* genes as described by Curtis & Meyers (2012). Approximately 500ng of template DNA was used in the PCR amplification of the 16S rRNA and *gyrB* genes, with 1µg of DNA being used for the amplification of the *atpD*, *recA*, *relA* and *rpoB* genes. The PCR products were purified using an MSB® Spin PCRapace kit (STRATEC Molecular, Germany) and sequenced (Macrogen, South Korea). Sequence analysis was performed using DNAMAN version 5.2.9 (LynnonBioSoft). MEGA version 5.05 (Tamura *et al.*, 2011) was used to conduct the phylogenetic analyses and to calculate the *gyrB* and concatenated-gene genetic distances using Kimura's 2-parameter model (Kimura, 1980). Phylogenetic trees were constructed using the maximum likelihood (Felsenstein, 1981), maximum parsimony (Takahashi & Nei, 2000) and neighbour-joining (Saitou & Nei, 1987) methods. Accession numbers for genes used in Figure2 are listed in Table S1.

Appendix 4K continued

All ISP media were prepared according to Shirling & Gottlieb (1966). Morphological properties were determined on ISP 2 and inorganic salts-starch agar (ISP 4). Physiological tests were Appendix K continued

carried out as described by Williams *et al.* (1989). All plates were incubated at 30°C, for the recommended periods, unless otherwise stated. Carbon source utilisation was tested as per the methods of Shirling & Gottlieb (1966), with all carbon sources being filter sterilised and tested at a final concentration of 1% (w/v), with the exception of the sodium salts, which were tested at 0.1% (w/v). Nitrogen source utilisation was performed as per Williams *et al.* (1989). All nitrogen sources were filter sterilised and tested at a final concentration of 0.1% (w/v). NaCl tolerance was determined on ISP 2 agar incubated for 14 days. Growth at pH 4.3, 7 and 10 and growth at different temperatures (20, 30 and 37°C) was determined on Bennett's agar (Atlas, 2004), incubated for 14 days.

Analysis of the isomer of diaminopimelic acid, the whole cell sugar pattern and the phospholipid pattern were carried out as described by Everest *et al*. (2013). Analysis of the respiratory quinones was carried out by the identification service at the DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen), Braunschweig, Germany. All chemotaxonomic analyses were performed on freeze dried cells of a culture of strain BC637^{T} grown in ISP 2 broth, with moderate shaking, at 30°C for 3 days. Fatty acid analysis was performed as a service by the BCCM/LMG Culture Collection, as per the recommendations of the commercial identification system MIDI (Microbial Identification System, Inc., Delaware, U.S.A.; MIDI Sherlock version 3.10; database: TSBA 50 rev 5.0), on cells grown at 28°C for 3 days on Tryptic Soy Agar (BBL 11768).

DDH analysis was performed as a service by the BCCM/LMG Culture Collection as described by Everest *et al*. (2013). Data are displayed as an average DNA-DNA hybridisation value with the difference between the means of the reciprocal values given in parentheses.

Appendix 4K continued

A query against the EzTaxon-e server (Kim *et al.*, 2012) with 1467 bp of 16S rRNA gene sequence indicated that strain BC637^T belonged to the genus *Kribbella* and was most closely related to *Kribbella flavida* DSM 17836^T (99.45% sequence similarity), *K. catacumbae* DSM 19601^T (99.32%), *K. albertanoniae*BC640^T (99.32%) and *Kribbella alba* YIM 31075^T(99.31%). However, the construction of a 16S rRNA gene maximum likelihood phylogenetic tree showed that strain BC637^T grouped with *Kribbella lupini* LU 14^T (99.17% sequence similarity), with low bootstrap support (60%) (Figure 1). Phylogenetic analyses based on the *gyrB-rpoB-recA-relA-atpD* concatenated gene sequences showed a strongly supported association of strain BC637^T with the type strains of *K. lupini* and *K. endophytica* (99.09% 16S rRNA gene sequence similarity; bootstrap value of 100% in the maximum likelihood tree, Figure 2).

Genetic distance values were calculated between strain BC637^T and all *Kribbella* type strains to assess the likelihood of this strain representing a novel genomic species.The *gyrB* based genetic distance values ranged from 0.016 to 0.096, with those between strain BC637^T and the type strains of *K. lupini, K. endophytica*, *K. flavida*, *K. catacumbae*, *K. albertanoniae* and *K. alba* being 0.026, 0.016, 0.051, 0.078, 0.078 and 0.094, respectively. The values for *K. lupini* (0.026) and *K. endophytica* (0.016) are below the 0.04 *gyrB* genetic distance threshold proposed to represent novel species in the genus (Kirby *et al.*, 2010). The concatenated five-gene genetic distance values ranged from 0.025 to 0.114. The values between strain BC637^T and the type strains of *K. lupini, K. endophytica*, *K. flavida*, *K. catacumbae*, *K. albertanoniae* and *K. alba* were 0.025, 0.025, 0.046, 0.059, 0.065and 0.070, respectively. The proposed threshold for this concatenated-gene sequence, above which strains can be assumed to belong to distinct genomic species, is 0.04 (Curtis & Meyers, 2012). As the concatenated-gene genetic distances between strain BC637^T and *K. endophytica* and between strain BC637^{T}and *K. lupini* were <0.04, DDH analysis was required to determine whether strain BC637^T belongs to a different genomic species.

DDH experiments revealed that strain BC637^T shared 63 (7) % DNA relatedness with *K. endophytica* DSM 23718^{T}and 63 (0) % DNA relatedness with *K. lupini* LU 14^T. Strain BC637^T

Appendix 4K continued

thus represents a separate genomic species when the threshold value of 70% DNA relatedness by DDH is used to delineate bacterial species (Wayne *et al.*, 1987). Although the DDH values are close to the 70% cut off, the MLSA results strongly suggest that BC637^T is a distinct genomic species: in all cases where there are multiple strains for a *Kribbella* species, the *gyrB* and concatenated-gene genetic distances between the type strain and all non-type strains of the same species are 0.000 (i.e. the *gyrB*, *rpoB*, *relA*, *recA* and *atpD* gene sequences for all strains of a species are identical). Strain BC637^T does not have any genetic distance values lower than 0.016 and 0.025 for the *gyrB* and concatenated-gene sequences, respectively, indicating sequence differences between strain BC637^T and all *Kribbella* type and non-type strains. Based on this evidence, strain BC637^T is not a strain of any of the established *Kribbella* species. The identical sequences between type strains and non-type strains are clearly shown in Figure 2 for the multi-strain species, *K. catacumbae, K. sancticallisti, Kribbella solani* and *Kribbella aluminosa*. The grouping of strain BC637^T with *K. endophytica* DSM 23718^T and *K. lupini* LU 14^T indicates a strong association between the strains, but the tree topology in Figure 2 and the genetic distance values suggest that the strains belong to distinct species. This pattern is also seen with the grouping of the closely related *Kribbella ginsengisoli* DSM 17941^{T}and*Kribbella koreensis* CIP 108301^T (concatenated-gene genetic distance value 0.015) and *K. solani* CIP 108508^T and *Kribbella hippodromi* S1.4^T (concatenated gene genetic distance value 0.017) in Figure 2. DDH analysis between strain BC637^T and the type strains of *K. flavida*, *K. catacumbae*, *K. albertanoniae* and *K. alba* was not performed, as the genetic distances between these strains were above the thresholds for distinguishing genomic species using both the *gyrB* and concatenated-gene sequences.

Physiological characterisation of strain BC637^T showed that it was phenotypically distinct from the type strains of *K. endophytica* and *K. lupini*. The results of the phenotypic characterisation of strain BC637^T are presented in Table 1 and in the species description. Chemotaxonomic analysis

Appendix 4K continued

of strain BC637^T showed that it had chemotaxonomic characteristics that are consistent with membership of the genus *Kribbella*: LL-DAP as the diagnostic diamino acid in the peptidoglycan; a phospholipid pattern characterised by the presence of phosphatidylcholine (and the absence of phosphatidylethanolamine); MK-9(H_4) as the major menaquinone; and ai-$C_{15:0}$ as the predominant fatty acid.

Appendix 4K continued

Description of *Kribbella italica* sp. nov.

Kribbella italica (i.ta'li.ca.N.L. fem. adj. *italica,* from Italy.)

Gram positive, catalase positive, oxidase positive, non-motile actinobacterium. Colonies appear convoluted with irregular edges on most media. Vegetative mycelium appears cream to white in colour. Hyphae are highly branched and fragment in both liquid and agar cultures. Aerial mycelium appears white on ISP 4. No diffusible pigment is produced on ISP 5. Melanin is not produced on peptone-yeast extract-iron agar (ISP 6) or tyrosine agar (ISP 7). Nitrate is reduced to nitrite. Produces H_2S. Utilises adonitol, L(+)-arabinose, D(+)-cellobiose, D(–)-fructose, D(+)-glucose,*myo*-inositol,inulin, α-lactose, D(+)-mannose, D(–)-mannitol, melibiose, raffinose, rhamnose, sucrose and D(+)-xylose as sole carbon sources, with weak growth on salicin and sodium acetate. Unable to utilise sodium citrate as a sole carbon source.Utilises DL-α-amino-n-butyric acid, L-arginine, L-asparagine, L-cysteine, L-histidine, L-4-hydroxyproline, L-methionine, potassium nitrate, L-serine, L-threonine and L-valine as sole nitrogen sources, with weak growth on L-phenylalanine. Grows at 20°C and 30°C, but not at 37°C. Grows optimally at pH 7, weakly at pH 10, but not at pH 4.3. Starch is hydrolysed. Casein, gelatin, hypoxanthine, Tween 80and L-tyrosine are degraded. Adenine is weakly degraded. Allantoin, urea, xanthine and xylan are not degraded. Grows weakly in the presence of 5% (w/v) NaCl. The cell wall peptidoglycan contains LL-DAP (Figure S2) and glycine (chemotype I (Lechevalier & Lechevalier, 1970)). Glucose, ribose and an unidentified pentose are present in the whole-cell sugar hydrolysate. The polar lipid profile includes diphosphatidylglycerol, phosphatidylcholine, phosphatidylglycerol, phosphatidylinositol, phosphatidylinositol mannoside, six unidentified aminolipids, four unidentified glycolipids, five unidentified phospholipids, an unidentified aminophospholipid and seven unidentified phosphoglycolipids (Figure S3). The major menaquinone of the type strain is MK-9(H_4) (75%), with minor amounts of MK-9 (4 %) and MK-9(H_2) (18 %).The major fatty acids (present in proportions >10% of total fatty acids) are ai-$C_{15:0}$ and $C_{17:1}\omega 8c$. The full fatty acid profile is shown in Table S2.

Appendix 4K continued

The type strain BC637^T (=DSM 28967^T = NRRL B-59155^T), was isolated from site CSC13 of the Saint Callixstus catacomb in Rome.

Acknowledgements

The authors are grateful to Dr Laura Bruno of Tor Vergata University in Rome and Dr R. Giuliani of the Pontificia Commissione di Archeologia sacra (C.d.V.) for their kind collaboration. G.J.E. held a Claude Leon Foundation Postdoctoral Fellowship. S.M.C. holds a National Research Foundation (NRF) Innovation Doctoral Scholarship. The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at are those of the authors and are not necessarily to be attributed to the NRF. This work was supported by research grants to P.R.M. from the NRF (grant number: 85476) and the University Research Committee (University of Cape Town).

Appendix 4K continued

References

Albertano, P., Moscone, D., Palleschi, G., Hermosı´n, B., Saiz-Jimenez, C., Sanchez-Moral, S., Hernandez Marine, M., Urzı`, C., Groth, I. & other authors (2003). Cyanobacteria attack rocks (CATS): control and preventive strategies to avoid damage caused by cyanobacteria and associated microorganisms in Roman hypogean monuments. In Molecular Biology and Cultural Heritage, pp. 151–162. Edited by C. Saiz-Jimenez. Lisse: Balkema.

Atlas, R. M. (2004).*Handbook of Microbiological Media (3rd edition)*. Boca Raton, Fl: CRC Press.

Cook, A. E. & Meyers, P. R. (2003). Rapid identification of filamentous actinomycetes to the genus level using genus-specific 16S-rRNA gene restriction fragment patterns. *Int J Syst Evol Microbiol* **53**, 1907-1915.

Curtis, S. M. & Meyers, P. R. (2012). Multilocus sequence analysis of the actinobacterial genus *Kribbella*. *Syst Appl Microbiol* **35**, 441-446.

Everest, G. J., Curtis, S. M., De Leo, F., Urzì, C. and Meyers, P. R. (2013). *Kribbella albertanoniae* sp. nov., isolated from a Roman catacomb, and emended description of the genus *Kribbella*. *International Journal of Systematic and Evolutionary Microbiology* **63** 3591-3596.

Everest, G. J. & Meyers, P. R. (2008). *Kribbella hippodromi* sp. nov., isolated from soil from a racecourse in South Africa. *Int J Syst Evol Microbiol* **58**, 443-446.

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *J Mol Evol* **17**, 368-376.

Kaewkla, O. & Franco, C. M. M. (2013). *Kribbella endophytica* sp. nov., an endophytic actinobacterium isolated from the surface-sterilized leaf of a native apricot tree. *Int J Syst Evol Microbiol* **63**, 1249-1253.

Kim, O. S., Cho, Y. J., Lee, K., Yoon, S. H., Kim, M., Na, H., Park, S. C., Jeon, Y. S., Lee, J. H., Yi, H., Won, S. & Chun, J. (2012). Introducing EzTaxon-e: a prokaryotic 16S rRNA Gene sequence database with phylotypes that represent uncultured species. *Int J Syst Evol Microbiol* **62**, 716–721.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16**, 111-120.

Kirby, B. M., Everest, G. J. & Meyers, P. R. (2010). Phylogenetic analysis of the genus *Kribbella* based on the *gyrB* gene: proposal of a *gyrB*-sequence threshold for species delineation in the genus *Kribbella*. *Antonie van Leeuwenhoek* **97**, 131-142.

Appendix 4K continued

Lechevalier, M. P. & Lechevalier, H. (1970). Chemical composition as a criterion in the classification of aerobic actinomycetes. *Int J Syst Bacteriol* **20**, 435-443.

Mohammadipanah, F., Hamedi, J., Göker, M., Fiebig, A., Pukall, R., Spröer, C. & Klenk, H.-P. (2013). *Kribbella shirazensis* sp. nov., isolated from Iranian soil. *Int J Syst Evol Microbiol* **63**, 3369-3374.

Park, Y. H., Yoon, J. H., Shin, Y. K., Suzuki, K., Kudo, T., Seino, A., Kim, H., Lee, J. S. & Lee, S. T. (1999). Classification of '*Nocardioides fulvus*' IFO 14399 and *Nocardioides* sp. ATCC 39419 in *Kribbella*gen. nov., as *Kribbella flavida* sp. nov.and *Kribbella sandramycini* sp. nov.*Int J Syst Bacteriol* **49**, 743-752.

Parte, A. C. (2014). List of prokaryotic names with standing in nomenclature.(http://www.bacterio.net/kribbella.html) Accessed May 2014.

Reasoner, D. J. & Geldreich, E. E. (1985). A new medium for the enumeration and subculture of bacteria from potable water. *Appl Environ Microbiol* **49**, 1-8.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406-425.

Shirling, E. B. & Gottlieb, D. (1966). Methods for characterization of *Streptomyces* species. *Int J Syst Bacteriol* **16**, 313-340.

Takahashi, K. & Nei, M. (2000). Efficiencies of fast algorithms of phylogenetic inference under the criteria of maximum parsimony, minimum evolution, and maximum likelihood when a large number of sequences are used. *Mol Biol Evol* **17**, 1251-1258.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony *Methods Mol Biol Evol* **28**, 2731-2739.

Trujillo, M. E., Kroppenstedt, R. M., Schumann, P. &Martínez-Molina, E. (2006).*Kribbella lupini* sp. nov., isolated from the roots of *Lupinus angustifolius*. *Int J Syst Evol Microbiol* **56**, 407-411.

Urzì, C. & De Leo, F. (2001). Sampling with adhesive tape strips: an easy and rapid method to monitor microbial colonization on monument surfaces. *J Microbiol Meth* **44**, 1**–**11.

Urzì, C., De Leo, F. & Schumann, P. (2008). *Kribbella catacumbae* sp. nov.and *Kribbella sancticallisti* sp. nov., isolated from whitish-grey patinas in the catacombs of St. Callistus in Rome, Italy. *Int J Syst Evol Microbiol* **58**, 2090-2097.

Appendix 4K continued

Wayne, L. G., Brenner, D. J., Colwell, R. R., Grimont, P. A. D., Kandler, O., Krichevsky, M. I., Moore, L. H., Moore, W. E. C., Murray, R. G. E., Stackebrandt, E., Starr, M. P. &Trüper, H. G. (1987). Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *Int J Syst Bacteriol* **37**, 463-464.

Williams, S. T., Goodfellow, M. & Alderson, G. (1989). Genus *Streptomyces* Waksman and Henrici 1943, 339^{AL}.*In Bergey's Manual of Systematic Bacteriology*, vol. 4, pp. 2452-2492.Edited by S. T. Williams, M. E. Sharpe & J. G. Holt. Baltimore: Williams & Wilkins.

Appendix 4K continued

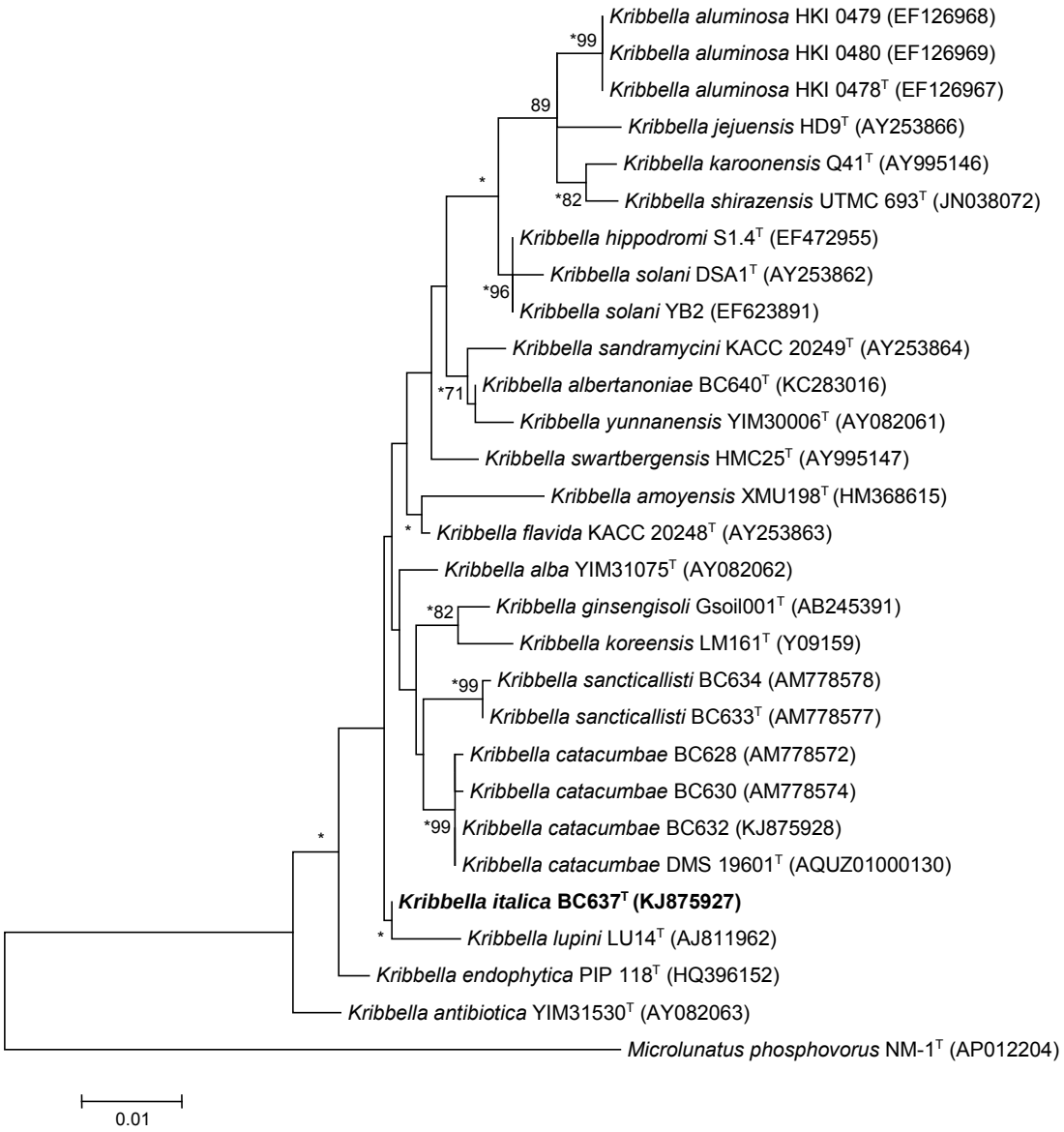
Figure legends

Figure 1.16S rRNA gene phylogenetic treeshowing the position of strain BC637^T within the genus *Kribbella.*The tree was constructed using the maximum likelihood method based on 1347 bp of sequence. Values at each node are the percentage bootstrap values of 1000 replications (only values ≥ 70% are shown), with asterisks (*) indicating the branches that were conserved in the maximum likelihood, neighbour-joining and maximum parsimony trees. Accession numbers are indicated in parenthesis after the strain numbers. The scale bar indicates 1 nucleotide substitution per 100 nucleotides. *Microlunatus phosphovorus* NM-1^T (AP012204) was used as an outgroup.

Figure 2.*gyrB-rpoB-recA-relA-atpD*phylogenetic treeshowing the position of strain BC637^T within the genus *Kribbella.*The tree was constructed using the maximum likelihood method based on 4099 bp of sequence. Values at each node are the percentage bootstrap values of 1000 replications (only values ≥ 70% are shown), with asterisks (*) indicating the branches that were conserved in the maximum likelihood, neighbour-joining and maximum parsimony trees. The scale bar indicates 2 nucleotide substitutions per 100 nucleotides. *Microlunatus phosphovorus* NM-1^T was used as an outgroup.Accession numbers of all the gene sequences used are listed in Table S1.

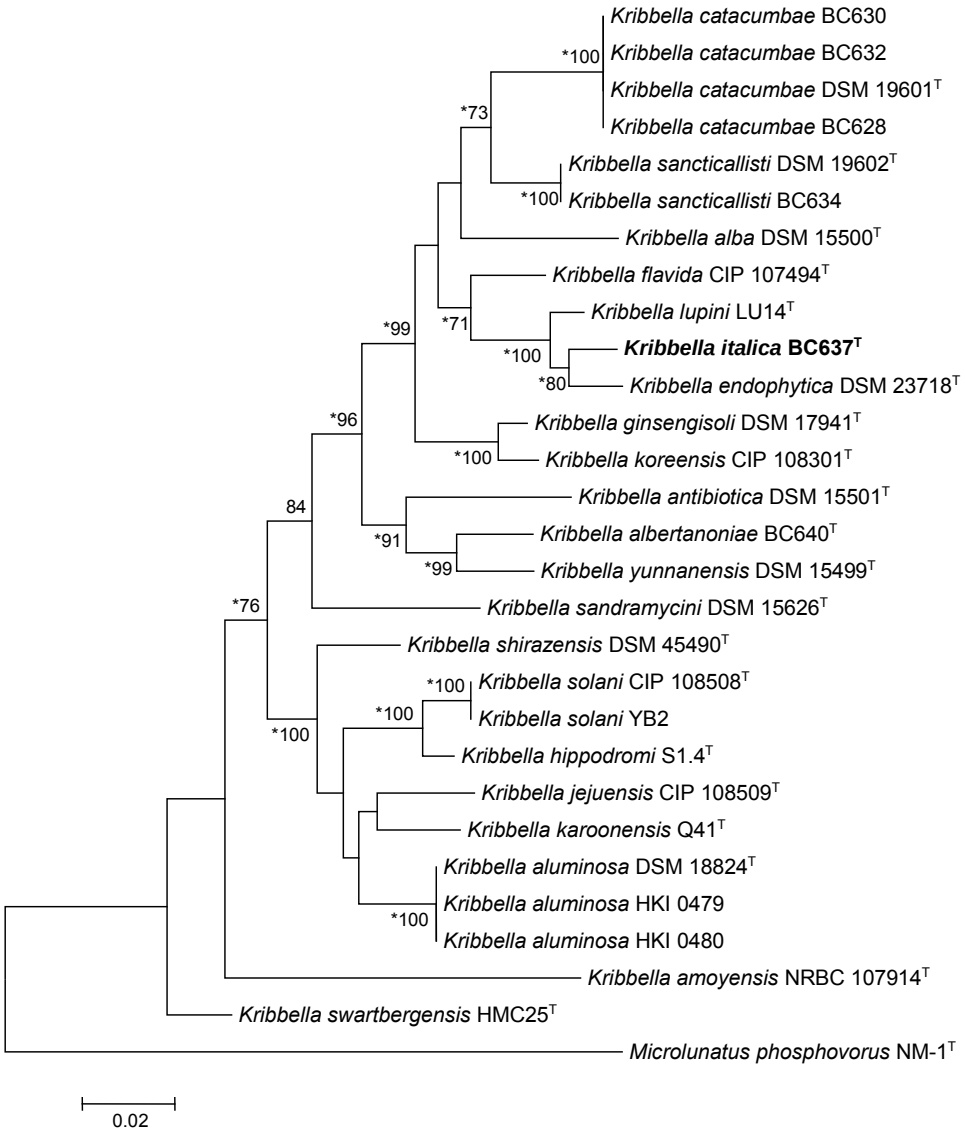
Appendix 4K continued

Figure 1 (16S rRNA)



Appendix 4K continued

Figure 2 (*gyrB-rpoB-recA-relA-atpD*)



Appendix 4K continued

Table 1. Phenotypic characteristics that allow differentiation of strain BC637^T from closely related *Kribbella* species.

Reference strains: 1, *Kribbella endophytica* DSM 23718^T; 2,*Kribbella lupini* LU 14^T. All data were determined in this study. Conflicting data (in parenthesis) are indicated. Symbols: ++, strong positive; +, positive; +w, weak positive;–, negative.

Test	**BC637^T**	**1**	**2**
Nitrate reduction	+	+w	+
Degradation of:			
Adenine	+w	+	–
Gelatin	+	+ (–[a])	+
Hypoxanthine	+	+ (–[a])	+
Urea	–	– (+[a])	–
Xanthine	–	+w (–[a])	+w
Xylan	–	–	+w (–[b])
Hydrolysis of starch	+	– (+[a])	–
Growth at:			
37°C	–	–	++
pH 4.3	–	+	+ (–[b])
5% (w/v) NaCl	+w	–	+
Utilisation as sole carbon source			
Inulin	++	+w	–
Raffinose	+	–	+
Salicin	+w	+w	+
Sodium acetate	+w	+w	–
Utilisation as sole nitrogen source			
L-cysteine	+	+w	+
L-methionine	+	+w	+w
Data determined by: [a], Kaewkla&Franko, 2012; [b],Trujillo*et al*. 2006.			

Appendix 4K continued

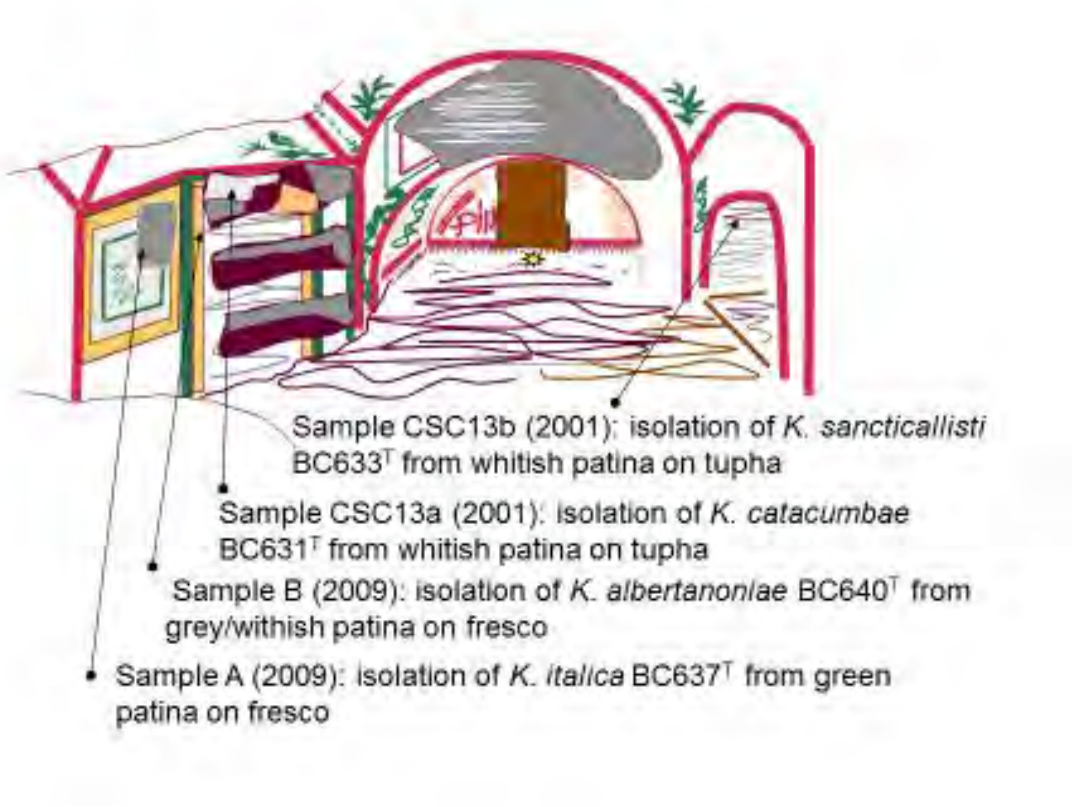


Figure S1: Diagram of the sampling sites in the Ocean cubicle (CSC13) in the Saint Callistus catacomb in Rome, Italy.

Appendix 4K continued

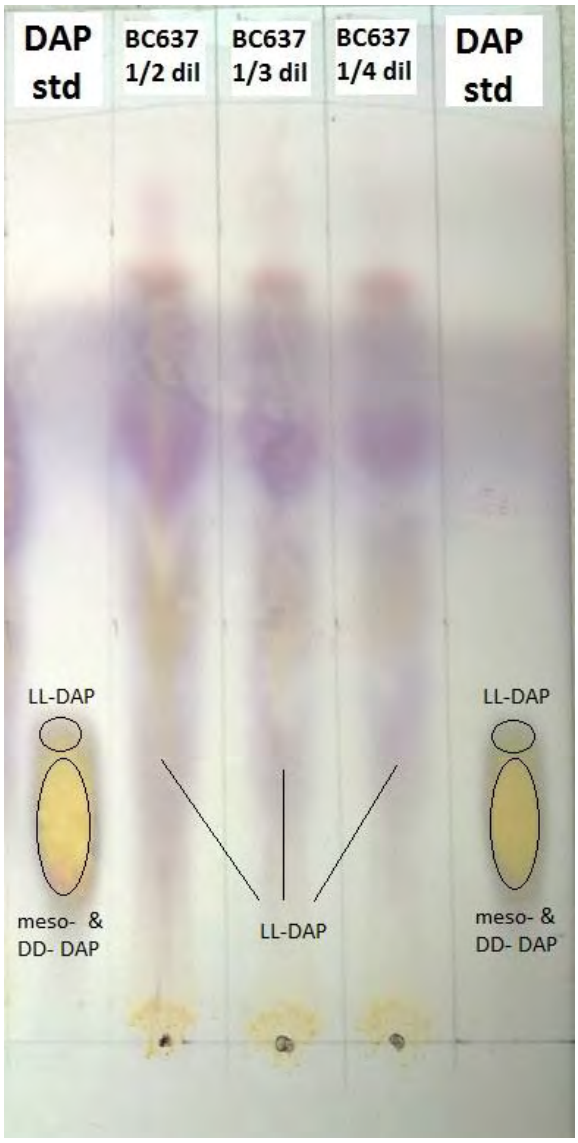


Figure S2: Identification of the diagnostic diamino acid in the peptidoglycan of strain BC637^{T}by thin layer chromatography. Abbreviation: DAP, diaminopimelic acid.

Appendix 4K continued

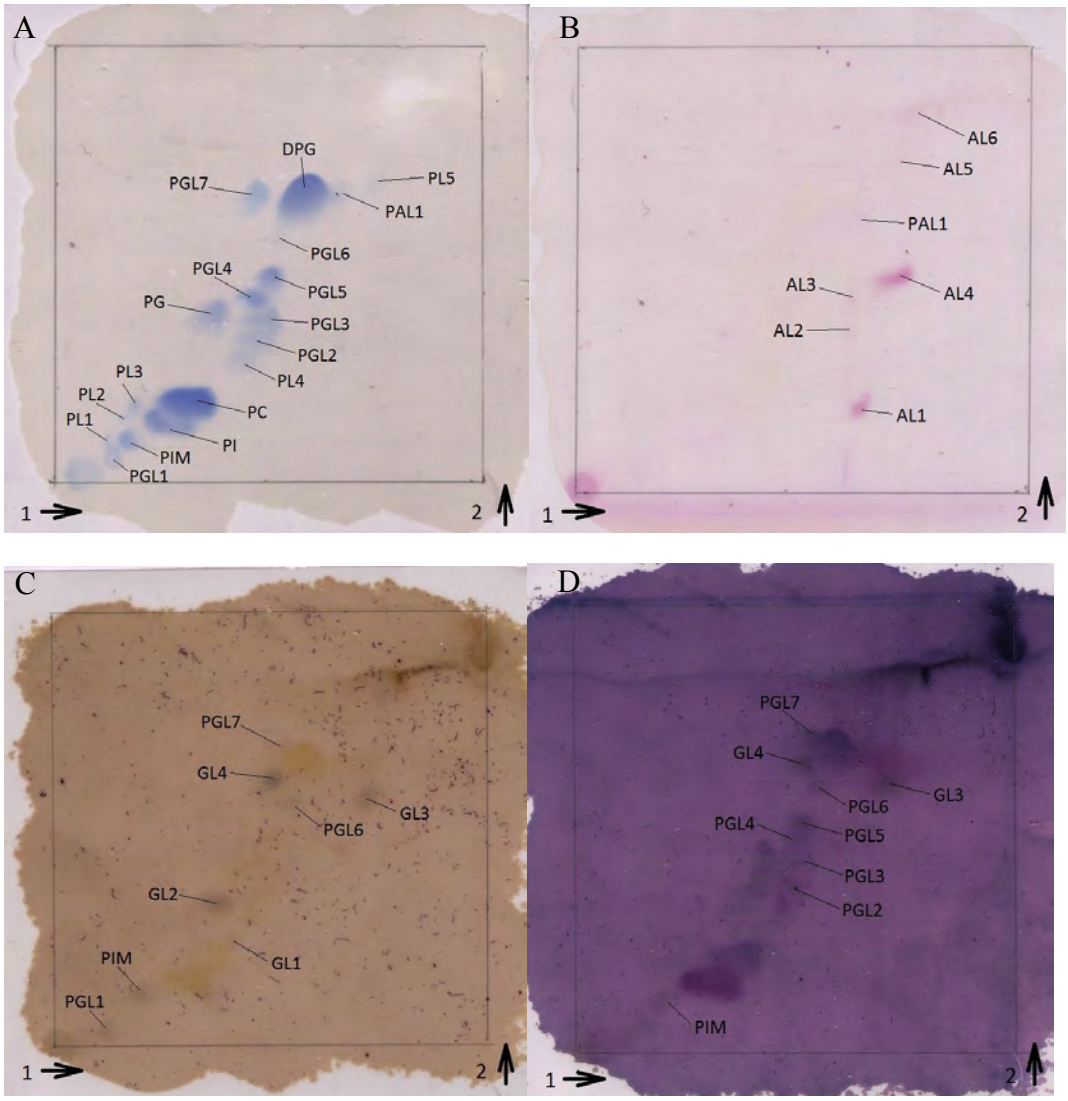


Figure S3: Polar lipid profile of strain BC637^T after two-dimensional thin layer chromatography and detection with (A) molybdenum blue, (B) ninhydrin, (C) α-naphthol and (D) *p*-anisaldehyde. Abbreviations: DPG, diphosphatidylglycerol; PC, phosphatidylcholine; PG, phosphatidylglycerol; PI, phosphatidylinositol; PIM,phosphatidylinositol mannoside; AL, unidentified aminolipids; GL, unidentified glycolipids;PAL, unidentified phosphoaminolipids; PGL, unidentified phosphoglycolipids; PL, unidentified phospholipids.

Appendix 4K continued

Table S1: Accession numbers of *Kribbella* genes used in this study.

Strain	*gyrB* gene accession numbers	*rpoB* gene accession numbers	*recA* gene accession numbers	*relA* gene accession numbers	*atpD* gene accession numbers
*Kribbella italica*BC637^T	KJ875917	KJ875919	KJ875921	KJ875923	KJ875925
Kribbella alba DSM 15500^T	EU434820	JX094181	JX094204	JX094227	JX094250
*Kribbellaalbertanoniae*BC640^T	KC283007	KC283009	KC283011	KC283013	KC283015
Kribbella aluminosa DSM 18824^T	EU434807	JX094182	JX094205	JX094228	JX094251
Kribbella aluminosa HKI 0479	GQ244403	JX094183	JX094206	JX094229	JX094252
Kribbella aluminosa HKI 0480	GQ244404	JX094184	JX094207	JX094230	JX094253
Kribbella amoyensis NBRC 107914^T	KC283006	KC283008	KC283010	KC283012	KC283014
Kribbella antibiotica DSM 15501^T	EU434819	JX094185	JX094208	JX094231	JX094254
Kribbella catacumbae DSM 19601^T	FJ917358	JX094186	JX094209	JX094232	JX094255
Kribbella catacumbae BC628	JX094273	JX094187	JX094210	JX094233	JX094256
Kribbella catacumbae BC630	JX094274	JX094188	JX094211	JX094234	JX094257
Kribbella catacumbae BC632	JX094275	JX094189	JX094212	JX094235	JX094258
Kribbella endophytica DSM 23718^T	KJ875918	KJ875920	KJ879522	KJ875924	KJ875926
Kribbella flavida CIP 107494^T	NC_013729	JX094190	JX094213	JX094236	JX094259
Kribbella ginsengisoli DSM 17941^T	JF775846	JX094191	JX094214	JX094237	JX094260
Kribbella hippodromi S1.4^T	EU434817	JX094192	JX094215	JX094238	JX094261
Kribbella jejuensis CIP 108509^T	EU434818	JX094193	JX094216	JX094239	JX094262
Kribbella karoonensis Q41^T	EU434816	JX094194	JX094217	JX094240	JX094263
Kribbella koreensis CIP 108301^T	EU434810	JX094195	JX094218	JX094241	JX094264
Kribbella lupini LU14^T	EU434811	JX094196	JX094219	JX094242	JX094265
Kribbella sancticallisti DSM 19602^T	FJ917357	JX094197	JX094220	JX094243	JX094266
Kribbella sancticallisti BC634	JX094276	JX094198	JX094221	JX094244	JX094267
Kribbella sandramycini DSM 15626^T	EU434812	JX094199	JX094222	JX094245	JX094268
Kribbella shirazensis DSM 45490^T	KF908048	KF908049	KF908050	KF908051	KF908052
Kribbella solani CIP 108508^T	EU434813	JX094200	JX094223	JX094246	JX094269
Kribbella solani YB2	EU434814	JX094201	JX094224	JX094247	JX094270
Kribbella swartbergensis HMC25^T	EU434808	JX094202	JX094225	JX094248	JX094271
Kribbella yunnanensis DSM 15499^T	EU434815	JX094203	JX094226	JX094249	JX094272

Appendix 4K continued

Table S2: Cellular fatty acids (%) of strain BC637^T. The cells were grown on Tryptic Soy Agar at 28°C for 3 days.

Fatty Acid	BC637^T
iso-$C_{13:0}$	0.55
iso-$C_{14:0}$	3.61
$C_{14:1}\omega5c$	0.40
$C_{14:0}$	1.07
iso-$C_{15:0}$	6.44
anteiso-$C_{15:0}$	35.01
$C_{15:1}\omega6c$	2.89
iso-$C_{16:1}$ H	2.06
iso-$C_{16:0}$	5.06
Sum in feature 3*	7.85
$C_{16:0}$	1.60
iso-$C_{17:1}\omega9c$	5.76
Sum in feature 4#	0.84
iso-$C_{17:0}$	3.34
anteiso-$C_{17:0}$	3.25
$C_{17:1}\omega8c$	14.49
$C_{17:1}\omega6c$	1.02
$C_{17:0}$	3.30
$C_{18:1}\omega9c$	1.47
*Sum in feature 3 contains $C_{16:1}\omega7c$ and/or iso-$C_{15:0}$ 2-OH #Sum in feature 4 contains iso-$C_{17:1}$ I and/or anteiso-$C_{17:1}$ B	

Chapter 5:
General Discussion

Chapter 5: General Discussion

The search for rare and novel actinomycetes in previously overlooked environmental niches has gained popularity in recent years as the need for novel natural products becomes urgent [Chaudhary *et al*., 2013]. Searching for endophytic actinomycetes that produce novel compounds has proven successful: recall the discovery of coronomycin, mentioned in Chapter 2, which is produced by an endophytic *Streptomyces* strain that was living in a vine in the Amazon jungle [Ezra *et al*., 2004]. Yet in order to enhance the search for these organisms and their compounds, a better understanding of their ecology and an improvement in taxonomy must be achieved [Kurtböke, 2012]. The standard use of 16S rRNA genes for identification and classification is insufficient at the species level and other gene-based approaches need to be considered [Santos and Ochman, 2004; Staley, 2006]. This study provides the tools to refine the search for new *Kribbella* species in a variety of natural environments, to identify if an isolate belongs to a new species in a quick and easy manner and to refine the phylogenetic analysis of the genus *Kribbella*, using genetic techniques.

The aims of this project were two-fold. The first aim was to develop methods to search for and isolate *Kribbella* strains from plant and soil samples. The second aim was to develop tools to enhance the phylogenetic analysis of the genus and to assess whether new isolates represent new species.

The first aim involved trying to isolate *Kribbella* species from soil samples. The plate isolations were successful in isolating 6 *Kribbella* colonies, which were identified using both 16S rRNA and *gyrB* gene sequences to be strains of *Kribbella solani*. In order to enhance the search for novel *Kribbella* strains, a *Kribbella* selective agar medium (KSM) was developed for use in the later isolation attempts and indigenous plant material was investigated for endophytic *Kribbella* strains. Plant tissue is home to a variety of endophytic species of bacteria. Here, 5 indigenous plants were used to try to isolate *Kribbella* strains from root and leaf material. The search for endophytic species of

Kribbella was motivated by the recently described endophytic species *K. endophytica* and *K. podocarpi* [Kaewkla and Franco, 2013; Curtis *et al*., manuscript under revision]. The KSM medium was developed in an effort to isolate rarer genera and deter the growth of the common genus *Streptomyces*. Although the KSM medium showed great promise in isolating *Kribbella* strains from a soil sample as part of an Honours student's project, the isolation of *Kribbella* colonies from plant material was not successful. This could be due to a number of reasons, including the screening of a very small amount of plant material (coupled with the low population density of endophytes), a rigorous surface sterilization protocol, and the fact that all of the plants were young, nursery-grown specimens planted in the same potting soil. The growth media for the isolations may be further improved upon by utilizing a lower concentration of nutrients to mimic the environmental conditions more closely, or adding plant extract to the media.

The latter part of this project involved metagenomically screening the environmental samples for the presence of *Kribbella* DNA. All the plant samples used for the isolation of *Kribbella* strains were also screened for the presence of *Kribbella* 16S rRNA gene sequences by PCR using a *Kribbella*-specific 16S rRNA gene primer. The Krb977F & R5 primer pair was used to amplify *Kribbella* sequences from the extracted environmental DNA, after which the PCR products were cloned, sequenced and identified using the Ez-Taxon database. The Krb977F primer was proven to be highly specific for *Kribbella* sequences, with only 4.7% of amplified sequences identified as belonging to non-*Kribbella* cultured strains. All five of the plant root and leaf samples were found to have *Kribbella* sequences present, even though none of these strains could be grown on agar plates. This *Kribbella*-specific detection tool has the ability to establish the presence of the target organism in environmental samples (soil, roots and leaves), before intensive plate isolations are used. The protocol may be used to target other genera of interest by simply applying other genus-specific primers to the sample.

Although no colonies of *Kribbella* were isolated on the KSM medium in this project, it was shown in an Honours study that the medium is selective for *Kribbella* species and would be beneficial to include in future *Kribbella* isolation attempts. However, the use of various media with different types and concentrations of nutrients and possibly plant extracts should be considered to increase the chances of isolating *Kribbella* strains.

A major part of this project was the development of the MLSA method for the genus *Kribbella* in order to improve and enhance the phylogenetic analysis. It was found previously that one housekeeping gene (*gyrB*) could improve upon the 16S rRNA gene sequence based phylogenetic analysis and thus four additional genes were chosen to be used in combination with the *gyrB* gene to enhance the phylogenetic analysis further. The genes chosen were required to meet certain guidelines that included, having a length of over 900 bp, being ubiquitous in bacteria, having a wide distibution on the genome and being present in a single copy per genome [Adékambi *et al*., 2011; Santos and Ochman, 2004; Zeigler, 2003]. The MLSA was an improvement on the standard 16S rRNA gene tree in that the branch lengths were longer (indicating higher resolution of strains) and had considerably higher bootstrap values (indicating greater confidence in the topology of the tree).

In the search for new species and new antibiotics, there is a strong need to assess if the organism is novel in a quick and easy manner. If the strain is novel, then further effort can be spent to fully characterize the strain and screen for antibiotic production. A quick method was developed and employed for the genus *Kribbella*. This involved using *gyrB* genetic distances to assess whether a strain is likely to belong to a new species and then further sequencing the genes *rpoB*, *recA, relA* and *atpD* to allow for an enhanced phylogenetic analysis of the strain in relation to the established species in the genus. This proved to be beneficial, as exemplified by the identification of the new species '*K. podocarpi', 'K. capetownensis', 'K. speibonae'* and '*K. italica'*.

In applying the method, genetic distance values are used to differentiate species. This provides a quick and easy way to determine if a strain is novel by using the *gyrB* gene sequence for the strain and comparing it to the *gyrB* gene sequences of the entire genus. The genetic distance values were calculated by using the Kimura 2 parameter method [Kimura, 1980]. It was found in previous work that a *gyrB* genetic distance of 0.04 was the threshold for determining if a species was novel [Kirby *et al*., 2010]. A genetic distance was also calculated using the concatenated genes (*gyrB*-*rpoB*-*recA*-*relA*-*atpD*) in which the threshold is also 0.04. However, it was found that the *gyrB* gene on its own is a good indicator of whether a strain is likely to represent a new species. Interestingly, it was found that in the multi-strain species (*K. aluminosa*, *K. catacumbae*, *K. sancticallisti* and *K. solani*), all strains of a species have a genetic distance of 0.000

with each other for both the *gyrB* and concatenated gene sequences (i.e. all strains of a species have identical sequences for each of the five genes).

Both the MLSA and genetic distance techniques were found useful in describing 4 new *Kribbella* species. The MLSA allowed for a more robust assessment of the phylogenetic neighbours of the new strains and provided higher resolution and more robust trees than those produced by using the 16S rRNA gene. The genetic distance was used to assess whether DDH was required to prove that the strain belonged to a new species. For strains YM53 (‘*K. capetownensis’*) and YM55 (‘*K. speibonae’*) the genetic distance values for the *gyrB* and concatenated genes were above the threshold of 0.04 and thus DDH was not needed to establish that these strains represent new species. In the case of strain BC637 (‘*K. italica’*), the *gyrB* and concatenated gene genetic distances to its closest relative, the type strain of *K. lupini*, were under the threshold values of 0.04, at 0.026 and 0.025, respectively. The genetic distance values for strain BC637 and the type strain of *K. endophytica* were also under the threshold at 0.016 and 0.025 for the *gyrB* and concatenated genes, respectively. The three strains were assessed by DDH and strain BC637 was found to share a 63% ± 0% DNA relatedness with the type strain *K. lupini* and a 63% ± 7% DNA relatedness to the type strain *K. endophytica*. As both of the DNA relatedness values are under the 70% similarity cut–off value for distinguishing between species [Wayne *et al*., 1987], strain BC637 was proven to belong to a new species [Everest *et al*., 2015]. The genetic distance values for strain YPL1 (‘*K. podocarpi’*) and the type strain of *K. karoonensis* were also under the genetic distance threshold values, at 0.032 for the *gyrB* gene and 0.038 for the concatenated gene. It was determined by DDH that these strains belong to separate species as they have a DNA relatedness value of 43% ± 13% [Curtis *et al*., manuscript under revision]. The examples of strains BC637, YPL1, YM55 and YM53 illustrate how using the MLSA and genetic distances enhances the description of new species. Strains YPL1 and BC637 illustrate the fact that, although they have concatenated genetic distance values to their closest relatives that are under the threshold for distinguishing species without requiring DDH, they are nevertheless distinct species. Other examples are, the type strains of *K. solani* and *K. hippodromi* (concatenated gene genetic distance 0.018, *gyrB* gene genetic distance 0.017, DNA relatedness 40.4 % ± 3.8%) [Everest and Meyers, 2008], *K. ginsengisoli* and *K. koreensis* (concatenated gene genetic distance 0.014, *gyrB* gene genetic distance 0.042, DNA relatedness 45% ± 5.37%) [Cui *et al*., 2010] and *K.*

albertanoniae and *K. yunnanensis* (concatenated gene genetic distance 0.033, *gyrB* gene genetic distance 0.052, DNA relatedness 53 % (± 8%) [Everest *et al*., 2013]. These examples suggest that the genetic distance threshold is too conservative and provide the motivation to propose a lower concatenated gene genetic distance threshold value for distinguishing between new species. Based on the data available, the threshold genetic distances for distinguishing between *Kribbella* species could be revised down to 0.02 for both the *gyrB* and concatenated-gene genetic distances.

Moving forward, it would be useful in the analysis of all new *Kribbella* isolates to apply the *gyrB* genetic distance method to assess quickly whether the strain is likely to belong to a novel species. If the *gyrB* gene genetic distance is >0.04, then the concatenated-gene genetic distances against all type strains of *Kribbella* could be calculated. Furthermore, it would be extremely useful to include an MLSA tree in all descriptions of new *Kribbella* species in order to keep the MLSA tree updated for the genus and to provide a more comprehensive phylogenetic analysis of the novel species being described.

Since the MLSA primers used in the study were designed from sequences in the family *Nocardioidaceae*, they are expected to amplify the same genes from all of the members of the family *Nocardioidaceae*. By amplifying and sequencing the genes *gyrB, rpoB, recA, relA and atpD* from all 119 type strains of the family *Nocardioidaceae*, the MLSA scheme could be applied to each of the eight other genera within the family (*Actinopolymorpha*, *Aeromicrobium*, *Flindersiella*, *Marmoricola*, *Mumia*, *Nocardioides*, *Pimelobacter* and *Thermasporomyces*). Moreover, a family level MLSA should improve and enhance the taxonomy of the family *Nocardioidaceae*, as well as each of the individual genera.

References

Adékambi, T., Butler, R. W., Hanrahan, F., Delcher, A. L., Drancourt, M. andShinnick, T. M., (2011). Core gene set as the basis of multilocus sequence analysis of the subclass Actinobacteridae. *PLoS ONE*. **6,** e14792

Chaudhary H.S., Soni, B., Shrivastava, A.R., Shrivastava, S. (2013). Diversity and Versatility of Actinomycetes and its Role in Antibiotic Production. *J. App. Pharm Sci.***3**, S83-S94

Cui, Y.-S., Lee, J.-S., Lee, S.-T. and Im, W.-T. (2010). *Kribbella ginsengisoli* sp. nov., isolated from soil of a ginseng field. *Int. J. Syst. Evol. Microbiol*. **60**, 364-368.

Curtis, S.M., Norton, I., Everest, G.J. and Meyers, P.R. (2015). *Kribbella podocarpi* sp. nov., isolated from the leaves of a yellowwood tree (*Podocarpus latifolius*) and *Kribbella capetownensis* sp. nov. and *Kribbella speibonae* sp. nov., isolated from soil. *Antonie van Leeuwenhoek.* Manuscript under revision.

Everest, G.J., Curtis, S.M., De Leo, F., Urzí, C. and Meyers, P.R. (2013). *Kribbella albertanoniae* sp. nov., isolatated from a Roman catacomb, and emended description of the genus *Kribbella*. *Int. J. Syst. Evol. Microbiol*. **63**, 3591-3596.

Everest, G.J., Curtis, S.M., De Leo, F., Urzí, C. and Meyers, P.R. (2015). Description of *Kribbella italica* sp. nov., isolated from a Roman catacomb. *Int. J. Syst. Evol. Microbiol.***64**: 000-000. *In Press*.

Everest, G.J. and Meyers, P.R., (2008). *Kribbella hippodromi* sp. nov., isolated from a soil from a racecourse in South Africa. *Int. J. Syst. Evol. Microbiol*. **58**, 443-446.

Ezra, D., Castillo, U.F., Strobel, G.A., Hess, W.M., Porter, H., Jensen, J.B., Condron, M.A.M., Teplow, D.B., Sears, J., Maranta, M., Hunter, M., Weber, B. and Yaver, D. (2004) Coronamycins, peptide antibiotics produced by a verticillate *Streptomyces* sp. (MSU-2110) endophytic on *Monstera* sp. *Microbiology.* **150**, 785-793.

Kaewkla, O. and Franco, C.M.M. (2013). *Kribbella endophytica* sp. nov., an endophytic actinobacterium isolated from the surface-sterilized leaf of a native apricot tree. *Int. J. Syst. Evol. Microbiol.***63**, 1249-1253.

Kimura M. (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol*. **16**, 111-120.

Kirby, B.M., Everest, G.J., and Meyers, P.R. (2010). Phylogenetic analysis of the genus *Kribbella* based on the *gyrB* gene: proposal of a *gyrB*-sequence threshold for species delineation in the genus *Kribbella*. *Antonie van Leeuwenhoek*. **97**, 131-142.

Kurtböke, D. Í. (2012). Biodiscovery from rare actinomycetes: an eco-taxonomical perspective. *Appl. Microbiol. Biotechnol*. **93**, 1843-1852.

Santos, S.R. and Ochman, H. (2004). Identification and phylogenetic sorting of bacterial lineages with universally conserved genes and proteins. *Environ. Microbiol*. **6**, 754-759.

Staley, J.T. (2006). The bacterial species dilemma and the genomic-phylogenetic species concept. *Phil. Trans. R. Soc. B***361**, 1899-1909.

Wayne, L.G., Brenner, D.J., Colwell, R.R., Grimont, P.A.D., Kandler, O., Krichevsky, M.I., Moore, L.H., Moore, W.E.C., Murray, R.G.E., Stackebrandt, E., Starr, M.P. and Trüper, H.G. (1987). International Committee on Systematic Bacteriology. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *Int. J. Syst. Bacteriol*. **37**, 463-464.

Zeigler, D. R. (2003). Gene sequences useful for predicting relatedness of whole genomes in bacteria. *Int. J. Syst. Evol. Microbiol*. **53**, 1893-1900.