Identification of endophytic bacterial communities associated with South African crops: *Sorghum bicolor* (L. Moench), *Pennisetum glaucum* and *Arachis villosulicarpa.*

Mapula Kgomotso Annah Maropola

A thesis submitted in partial fulfilment of the requirements for the degree of MAGISTER SCIENTIAE (M.Sc.)

Department of Biotechnology

University of the Western Cape

Bellville

Supervisor: Prof. Marla Tuffin

Co-Supervisor: Dr. Jean-Baptiste Ramond

20 May 2014

DECLARATION

I, Mapula Kgomotso Annah Maropola, hereby declare that *Identification of endophytic bacterial communities associated with South African crops: Sorghum bicolor (L. Moench), Pennisetum glaucum and Arachis villosulicarpa* is my own work, that it has not been submitted for any degree or examination in any other universty, and that all the sources I have used or quoted have been indicated and acknowledged by complete.

Date: .... .. May 2014

Signed:...
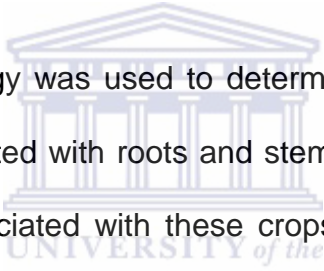
ABSTRACT

Endophytic bacteria live in plant tissues, and are constantly interacting with the host plant. These interactions could be beneficial to the plant, whereby the bacteria promote plant growth or enhance the plant's resistance to disease and environmental stress; or they could be detrimental to plant life when parasitic or pathogenic bacteria are involved. In this study, the diversity of endophytic bacteria associated with food crops, sorghum (*Sorghum bicolor* L. Moench), pearl millet (*Pennisetum glaucum* L.) and groundnut (*Arachis villosulicarpa*) is investigated using culture-independent techniques: terminal retriction fragment length polymorphism (t-RFLP) and next generation sequencing (NGS).
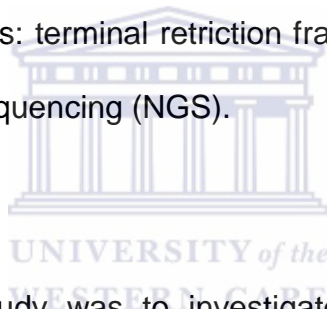
The first objective of this study was to investigate the effect of different DNA extraction protocols on mDNA yield and quality, as well as the diversity of endophytic bacteria retrieved from root and stem tissues (0.1g or 0.3g) of sorghum, pearl millet and groundnut. Protocols used include two classical methods (CTAB- and SDS-based) and five commercial kits: MoBio PowerPlant Pro® DNA Isolation Kit, Qiagen DNeasy^R Plant Mini Kit, Fermentas GeneJET Plant Genomic DNA Purification Kit, MoBio PowerSoil^{TM} DNA Purification Kit and MoBio UltraClean® Soil DNA Isolation Kit. Eletrophoresis and the Nanodrop were used to determine DNA yield and purity. The quality of mDNA was further analysed in PCR-amplification of the bacterial 16S rRNA gene. T-RFLP was used to determine the diversity of communities retrieved with the different methods. Classical protocols were shown to retrieve the highest mDNA yield from all tissues compared to commercial kits; however, the quality of

mDNA was compromised, particularly groundnut mDNA. Commercial kits were more consistent in retrieving mDNA of good PCR quality; however, they underestimated the diversity of endophytic bacteria the most. The SDS-protocol was shown to retrieved the most diverse endophytic communities from monocotyledonous plants at a higher starting plant material (0.3g). The CTAB protocol was the most efficient process to use on groundnut tissues; however, this process needs to be further optimised. This study emphasizes the need to continuously evaluate routine laboratory techniques in order to limit process-introduced biases in metagenomic studies of endophytic communities.

454 pyrosequencing technology was used to determine the diversity of endophytic bacterial communities associated with roots and stems of sorghum and pearl millet. Endophytic communities associated with these crops are diverse. Dominant phyla included Actinobacteria, Bacteroidetes, Firmicutes and Proteobacteria. Dominant bacterial genera found in both plants such as *Paenibacillus*, *Agrobacterium*, *Pseudomonas* and *Erwinia* are known to have a diverse range of metabolic capabilities, and can be targeted for production agricultural bio-inoculants (biofertilizers and biocontrol), bioremediation and many other industrial applications. Certain genera appeared to be plant-species specific, including the *Sphingobium*, *Sphingomonas*, *Rhizobium*, *Herbaspirillum* that were only dominant in sorghum tissues, and the *Arthrobacter*, *Chryseobacterium* and *Exiguobacterium* found in pearl millet tissues. This study shows that the ecology of sorghum and pearl millet endophytic communities needs to be further explored in order to understand their role in plant health and growth.

ACKNOWLEDGEMENTS

DEDICATION

To my family, who bore my long absences: Mama, Papa, Kabelo, Tebogo, Kgabo, Mathopa, Keabetswe, Tekano and Kgasago. This is also for my wonderful friends that have helped me see my dream through, even when I lost sight of it: Leshoto, Kenalemang, Amelie, Rinky, Timothy, Godfrey, Gerda, Sharon, and many others.
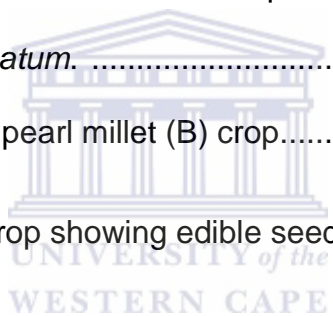
LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

UNIVERSITY *of the*
WESTERN CAPE

LIST OF TABLES

Chapter 3

Chapter 4

Chapter 5

LIST OF ABBREVIATIONS

α	alpha
ABA	abiscisic acid
ACC	1-aminocyclopropane-1-carboxylate
ANOVA	analysis of variance
APS	adenosine phosphosulphate
ATP	adenosine triphosphate
β	beta
γ	gamma
^{o}C	degrees celsius
bp	base pairs
CTAB	hexadecyltrimethylammonium bromide
CP	classical protocol
CPmDNA	metagenomic DNA retrieved with classical protocols
CK	commercial kit
CK_p	commercial kit designed for plant DNA extraction
CK_s	commercial kit designed for soil DNA extraction
CKmDNA	metagenomic DNA retrieved with commercial kits
DGGE	denaturing gradient gel electrophoresis
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
EDTA	ethylene diamine tetra acetic acid
et al	et alia
FAM	fluorescein amidite dye
FAME	fatty acid methyl esters
FISH	Fluorescent *in situ* hybridisation

g	grams
GAs	Giberellins
H'	Shannon diversity index
HCl	Hydrochloric acid
IAA	indole acetic acid
kbp	kilobase pairs
kcal	kilocalories
KCl	potassium chloride
KH_2PO_4	potassium di-hydrophosphate
kg/ha	kilograms per hectares
μL	microliter
μm	micrometer
$μg.mL^{-1}$	microgram per milliliter
M	molar
mDNA	metagenomic DNA
mg	milligrams
$mg.mL^{-1}$	milligrams per milliliter
min	minute
mL	milliliter
mM	millimolar
mm	millimeter
nMDS	non-metric multidimensional plot
ng	nanogram
N/ha	nitrogen per hectare
NaCl	sodium chloride
$Na_2HPO4.2H_2O$	sodium orthophosphate dihydrate

NBRIP medium	National Botanical Research Institute's phosphate
NGS	Next Generation Sequencing
OTU	operational taxonomic unit
PBS	phosphate buffer solution
PCR	polymerase chain reaction
PGPMs	plant growth promoting microorganisms
PGPBs	plant growth promoting bacteria
PGPEBs	plant growth promoting endophytic bacteria
PLFA	phospholipid fatty acid
PPi	pyrophosphate
PSMs	phosphate solubilising microorganisms
PVK	Pikovskaya medium
PVPP	polyvinylpyrrolidone
rpm	revolutions per minute
rcf	relative centrifugal force
SDS	Sodium dodecyl sulphate
SO_4^{2-}	sulphate ion
TAE	tris-acetic acid EDTA
TE	tris-EDTA
ton/ha	tonnes per hectare
t-RFs	terminal restriction fragments
t-RFLP	terminal restriction fragment length polymorphism
Tris	tris (hydroxymethyl)-aminomethane
U/μL	units per microliter
V	volts
v/v	volume per volume

CHAPTER 1

Literature Review

CONTENTS

1.1. Overview of plant-microbe relationships

Microbial life is ubiquitous in all environments. As such, plants are constantly interacting with various microorganisms throughout their development (Hallmann *et al*., 1997). These interactions take place in the different plant-created microenvironments, which also provide distinctive habitats for microbial colonisation (Morgan *et al*., 2005). Plant-microbe interactions play an important role in the host plant's development and health. Mutualistic relationships involve microorganisms that promote plant growth and/or enhance the plant's resistance to diseases (Hirsh, 2004). These beneficial microorganisms are thus referred to as plant growth promoting microorganisms (PGPMs). Commensalistic microorganisms have no impact on the plant's development whereas parasitic and pathogenic microorganisms reduce the plant's fitness and cause disease (Hirsh, 2004; Morgan *et al*., 2005). These interactions also indicate the feeding patterns of the microorganisms, and are therefore referred to as trophism states (Newton *et al*., 2010).

This review thus discusses, broadly, the plant-microbe interactions in the different plant-influenced microenvironments. The current study is on the diversity of endophytic bacteria associated with agricultural crops (sorghum and pearl millet); therefore, the establishment of endophytic communities will be discussed in greater detail, including their significance to plant life and the broader biotechnology field. Different methodological approaches (culture-based and culture-independent) that

are currently used in the study of plant-associated endophytic bacterial communities are also discussed.

1.1.1. Diversity of plant-associated microorganisms

Plants interact with a broad range of microorganisms. The most studied of these are fungi (Strobel and Daisy, 2003). Fungal representatives include yeasts and filamentous fungi found in phyla such as Ascomycota and Basidiomycota, as well as subphyla Mucoromycotina (Buée *et al.*, 2009). The interactions between mycorrhizal fungi (ectomycorrhizal and arbuscular mycorrhizal fungi) have been well studied due to the various contributions of these fungi to agricultural and forest soil ecology (Buée *et al.*, 2009; Raaijmakers *et al.*, 2009). Some fungal groups are well-known plant-pathogens. A few *Fusarium* species (e.g. *F.avenaceum*, *F. oxysporum, F. graminearum, F. culmorum*) are responsible for crop diseases such as crown rot, head blight and wilt in agricultural crops including barley, maize, wheat and sorghum; and they have caused great economic losses worldwide (Stępień and Chełkowski, 2010; Tunali *et al.*, 2012).

Bacteria constitute the second most studied group of plant-associated microorganisms. Bacterial taxa including Proteobacteria, Firmicutes, Cyanobacteria, Cytophaga-Flavobacterium-Bacteriodetes and Actinobacteria have been isolated from plant tissues and rhizospheric soils (Rosenblueth and Martinez-Romero, 2006). As with fungi, bacterial interactions with the plants can be beneficial or detrimental to

plant life, and these associations will be discussed in greater detail in subsequent paragraphs.

Archaea, phages, protozoa, algae and microathropods have also been found in the plant-influenced environments, but have been studied in much less detail (Raaijmakers *et al.*, 2009).

1.1.2. Plant-created microenvironments

Plant-created microenvironments include the rhizosphere, phyllosphere and endosphere (Morgan *et al.*, 2005; Montesinos, 2003).
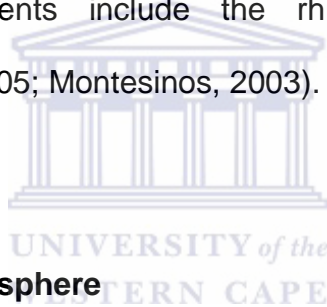
1.1.2.1. The rhizosphere

The bulk of soil adjacent to and is influenced by the plant roots is known as the rhizosphere (Morgan *et al.*, 2005). The root exudates released into this environment are important for the plant's development and health as they serve as mineral ion chelators, plant-growth promoting phytohormones (e.g. gibberellins, auxins and indole acetic acid), immune response phytochemicals (e.g. salisylic acid, jasmonic acid and ethylene), and biocatalytic enzymes (Faure *et al.*, 2008). Root-secreted mucilage promotes soil particle aggregation, thus increasing the water retention potential of the soil (Walker *et al.*, 2003). Root exudates also provide nutrition for soil microorganisms as they contain biomolecules such as sugars, fatty acids, nucleotides, organic acids, phenolics, plant growth regulators, putrescine, sterols

and vitamins (Lugtenberg and Kamilova, 2009). As a result, microorganisms are recruited from the surrounding bulk soil into the rhizosphere, where complex plant-microbe interactions are established (Compant *et al.*, 2010) (Berg and Smalla, 2009).

a. Structure of the rhizosphere

The rhizosphere is separated into four ecological niches (Figure 1.1) (McNear and David, 2013; Morgan *et al.*, 2005). The ectorrhizosphere is the soil environment that is immediately adjacent to the root. The interface between the soil matrix and the root surface constitutes the rhizoplane, and the endorhizosphere is the root tissue itself (Morgan *et al.*, 2005). Mycorrhizal fungi that associate with the roots of certain plants form an extensive hyphal network around plant roots, which extends beyond the ectorrhizosphere. The entire area covered by these hyphae is known as the mycorrhizosphere, and it is important because the mycorrhizal hyphae increase the plant's access to moisture and nutrients further from the root (Morgan *et al.*, 2005).
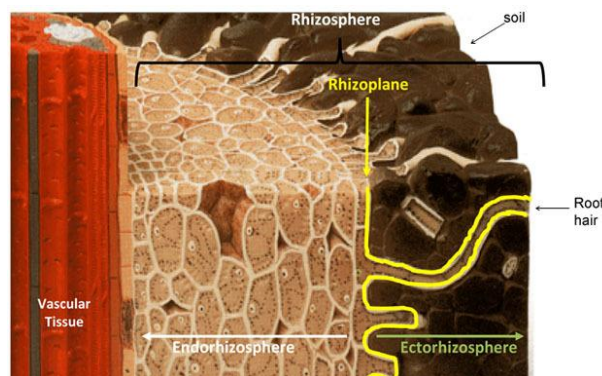
Figure 1.1. Schematic of a root section showing the structure of the rhizosphere (Adapted from McNear and David, 2013).

b. <u>Microbe recruitment into the rhizosphere</u>

The rhizophere is characterised by increased microbial activity, microbial species richness and diversity compared to the surrounding bulk soil (Compant *et al.*, 2010), and this is due to the high-nutrient content of rhizospheric soils. The spatial shift in microbial composition between bulk soils and rhizospheric soils was highlighted in a community profiling study using phospholipid fatty acid (PLFA) assay, whereby rhizospheric bacterial communities of jatropha and switchgrass were shown to be more abundant and diverse than bulk soil communities (Chaudhari *et al.*, 2012).

Recruitment of microorganisms into the rhizosphere is a selective process. Microbial surface receptors recognise specific chemicals (chemoattractants) in the root exudates, and this triggers a chemotactic response towards the roots. This movement is facilitated by bacterial flagella or growth of fungal hyphae (Lugtenberg and Kamilova, 2009; Compant *et al.*, 2010). The presence of specific chemoattractants in the rhizospheric soils is dependent on the root exudate composition, which in turn is determined by plant species and its developmental stage as well as biotic (e.g. presence of pathogens) and abiotic stresses (e.g. drought conditions) (Wieland *et al.*, 2001; Griffiths *et al.*, 2003; Berg and Smalla, 2009).

Colonisation of specific micro-niches in the rhizosphere is dependent on the nutritional requirements of the microorganisms and suitability of physical conditions for optimal growth (Compant *et al.*, 2010). For example, rhizoplane-colonising

Pseudomonas fluorescens (WCS365) establish microcolonies in the crevices of the host plant's root surface (Figure 1.2), where they secrete mucigel to form a protective sheath (Chin-A-Woeng *et al.*, 1997).
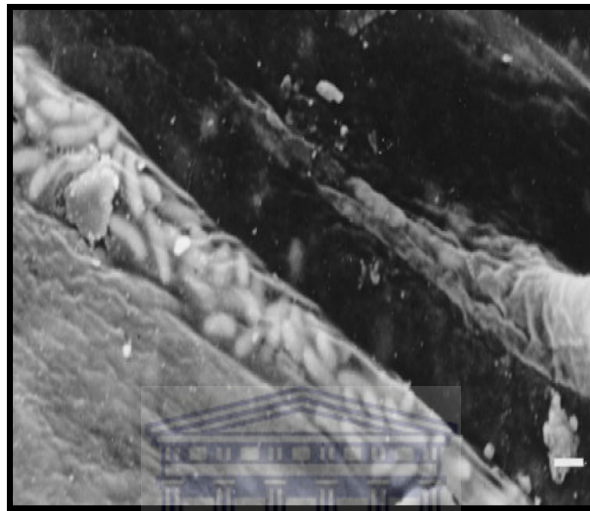
Figure 1.2. Rhizoplane colonisation. Part of a mucigel-protected *Pseudomonas flourescens* (WCS365) microcolony formed in the epidermal-cell junction of a 3-day old tomato root viewed using scanning electron microscopy (bar represents 1μm) (Adapted from Chin-A-Woeng *et al.*, 1997).

Rhizospheric microbial communities are also shaped by edaphic factors. Soil pH has been singled out as a key deterministic factor in the localisation of specific bacterial groups across different ecosystems (Fierer and Jackson 2006) as well as the composition and diversity of rhizospheric bacterial communities (Ramond *et al.* 2013). Water availability was shown to only affect rhizospheric communities that are seldom exposed to moisture stress in their natural environment (Fierer *et al.*, 2003).

1.1.2.2. The phyllosphere

The surfaces of the aerial parts of the plant constitute a microenvironment called the phyllosphere. This is the most nutrient-deficient plant-influenced microbial habitat due to very little exudates released by the plant (Whipps *et al.*, 2008). It is also exposed to environmental factors such as wind and rain (Kroupitski *et al.*, 2011). Nutrient deficiency and the impermeable, water resistant cuticle covering leaf surfaces create a barrier for microbial colonisation (Whipps *et al.*, 2008). However, microorganisms are transferred to the plant surface from the atmosphere or by vectors such as animals and insects. Competent phyllosphere colonisers are termed epiphytes (Whipps *et al.*, 2008).
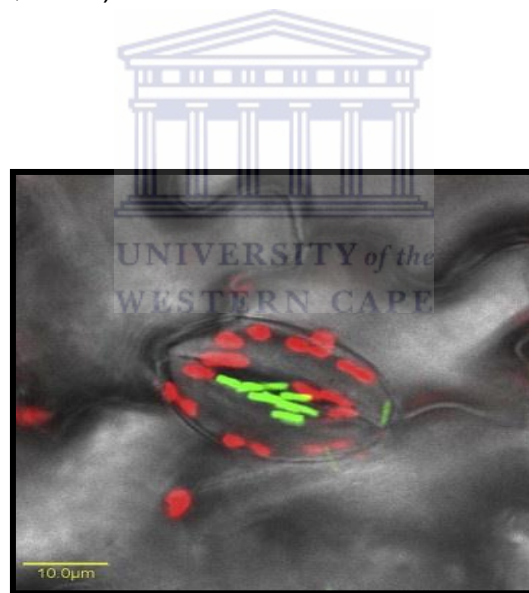
Figure 1.3. Colonisation of a romaine lettuce leaf stomatal opening (guard cell mitochondria are stained with a red dye) by *Salmonella typhimurium* (tagged with green fluorescent protein) viewed with a confocal microscope (Adapted from Kroupitski *et al.*, 2011).

Microorganisms present on the plant surface, but are unable to colonise the phyllosphere are known as transient epiphytes. Residual (true) epiphytes are able to proliferate and establish communities in this environment. True epiphytes form colonies in areas where there is little release of plant metabolites and protection from environmental factors (Whipps *et al.*, 2008; Kroupitski *et al.*, 2011) such as stomatal openings (Figure 1.3), depressions in the cuticle, along the leaf veins, at the base of the trichomes, near the hydrathodes and on pectate hairs (Whipps *et al.*, 2008).

1.1.2.3.　The endosphere

The endosphere is the internal environment of the plant (Wieland *et al.*, 2011). Endophytic microorganisms are thus defined as microorganisms found within surface-sterilised plant organs (Hardoim *et al.*, 2008). They are recruited from the surrounding environment through the roots or aerial parts of the plant in a process known as lateral transmission, or they can be passed from generation to generation of host plants through seeds or vegetative tissues in a process called vertical transmission (Hardoim *et al.*, 2008). Microorganisms are considered to be competent endophytes when they are able to infect plant tissues *and* also survive and proliferate within this environment. Two types of competent endophytes are recognised: obligate endophytes that spend their whole life cycle *in planta* and facultative endophytes that spend only a portion of their life cycle within plant tissues (Hardoim *et al.*, 2008).

Endophytic microorganisms grow in the intercellular spaces in plant tissues. Bacteria form microcolonies and fungal hyphae grow between the plant cells and branch out across the plant tissue (Rosenblueth and Martínez-Romero, 2006). The extracellular fluid released by plant cells contains photosynthetic products, phytohormones, enzymes and cellular metabolites, and thus provides nutrition for endophytic microorganisms (Cutler *et al.*, 2007). As with rhizospheric and phyllosphere communities, endophytic communities are shaped by various biotic and abiotic factors including plant species type, tissue type, plant age, seasonal changes and soil type (Kuklinsky-Sobral, 2004; Conn and Franco, 2004).

Endophytes have been isolated from all plant organs including roots, stems, leaves, flowering and fruiting bodies as well as seeds (Aravind *et al.*, 2009; Compant *et al.*, 2011; Fürnkrans *et al.*, 2012). Different plant organs create unique and complex endophytic microenvironments as a result of their structural and functional differences. The main plant organs are discussed below.

(a) The roots

The main functions of the roots are to anchor the plant in the soil and to absorb water and essential nutrients (Hopkins *et al.*, 2004). Roots of dicotyledonous plants are characterised by a tap-root structure. This consists of lateral roots developing from a primary root that is a below-ground extension of the shoot. Monocotyledonous roots have an adventitious structure that lacks primary root, and grow directly from the shoot (Hopkins *et al.*, 2004). Figure 1.4A shows the different zones of

development of a typical root. Root tips are covered by the protective root cap. Above the root tip is the region of cell division where root growth takes place (Cutler *et al.*, 2007). Regions of cell elongation and cell differentiation are characterised by cell development and specialization (Hopkins *et al.*, 2004). In dicotyledonous plants, lateral roots emanate from the region of cell differentiation. The arrangement of specialised plant tissues in the root organ are illustrated in Figure 1.4B.
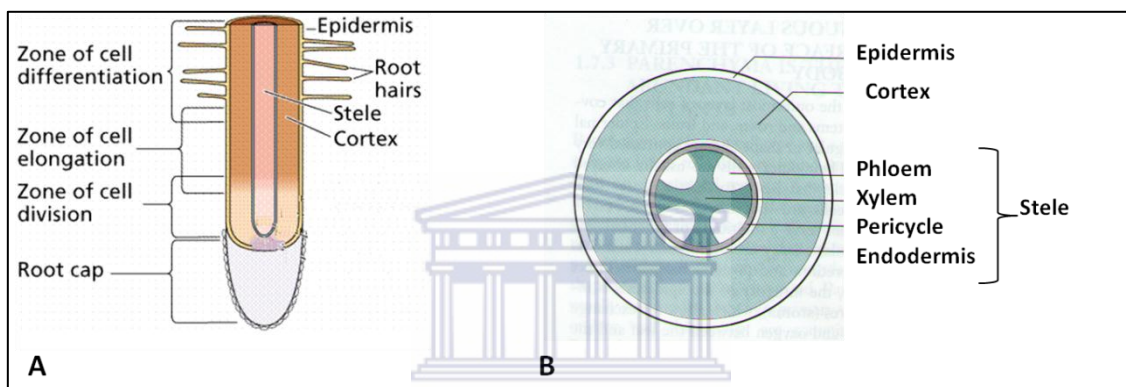
Figure 1.4. Schematic diagram of (A) a longitudinal section and (B) a cross section of a typical plant root (Adapted from Hopkins *et al.*, 2004).

The outermost tissue of the root is the protective epidermis. Adjacent to the epidermis is the cortical tissue, made up of parenchymal cells. These store nutrients for the roots, and allow movement of water and nutrients between the external environment and the vascular system (Hopkins *et al.*, 2004). The stele is made up of the endodermis, pericycle, xylem and phloem. The endodermis separates the cortex from the vascular system and the pericycle is an area of lateral root development (Figure 1.4B). The rigid lignin-rich xylem is responsible for transportation of water and solutes from the soil to aerial parts of the plant, and it also maintains the

structural integrity of the plant. Phloem transports photosynthetic products from the aerial parts of the plants to the roots (Hopkins *et al.*, 2004).

(b) <u>The stem</u>

The plant stem grows upward from the shoot. It branches to support and position the leaves for efficient exposure to sun radiation needed for photosynthesis (Hopkins *et al.*, 2004). Stem tissue arrangement is similar to arrangement in roots, with the exception that in stems, vascular tissues are arranged in "bundles" instead of the stele. In dicotyledonous stems, the vascular bundles form a ring in the ground tissue whereas in monocotyledons, they are "scattered" in the ground tissue (Figure 1.5). The epidermis of herbaceous stem is covered by a thin cuticle layer to prevent desiccation and microbial colonisation, and in woody plants, it is reinforced with lignin to form a protective bark (Cutler *et al.*, 2007).
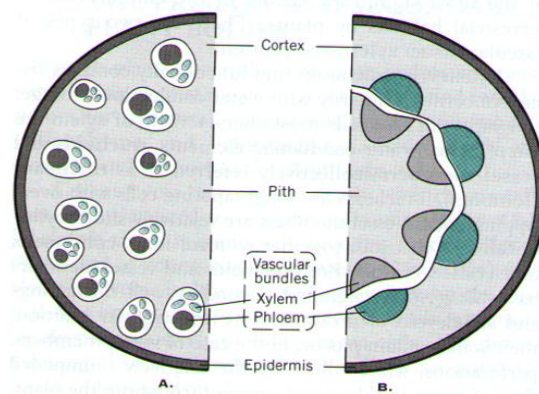
Figure 1.5. Schematic diagram of a cross-section of (A) a monocotyledonous stem and (B) a dicotyledonous stem (Adapted from Hopkins *et al.*, 2004).

(c) <u>The leaves</u>

The leaves are the photosynthetic organs of the plant. The chloroplast-containing mesophyll tissue captures sun radiation and converts carbon dioxide and water to complex energy-rich carbon compounds (Cutler *et al.*, 2007). Figure 1.6 is an illustration of the tissue arrangement in a dicotyledonous leaf.

Leaf blades are thin and covered with a water-resistant cuticle layer. Stomatal openings on the leaf surface allow gas exchange between the atmosphere and the leaf tissues. The vascular system in leaves is organised in leaf veins that facilitate transportation of water and nutrients, as well provide structural support for the leaf blade (Figure 1.6). Leaf vein endings at the leaf edges are called hydathodes, and these release little moisture and solutes (Cutler *et al.*, 2007).
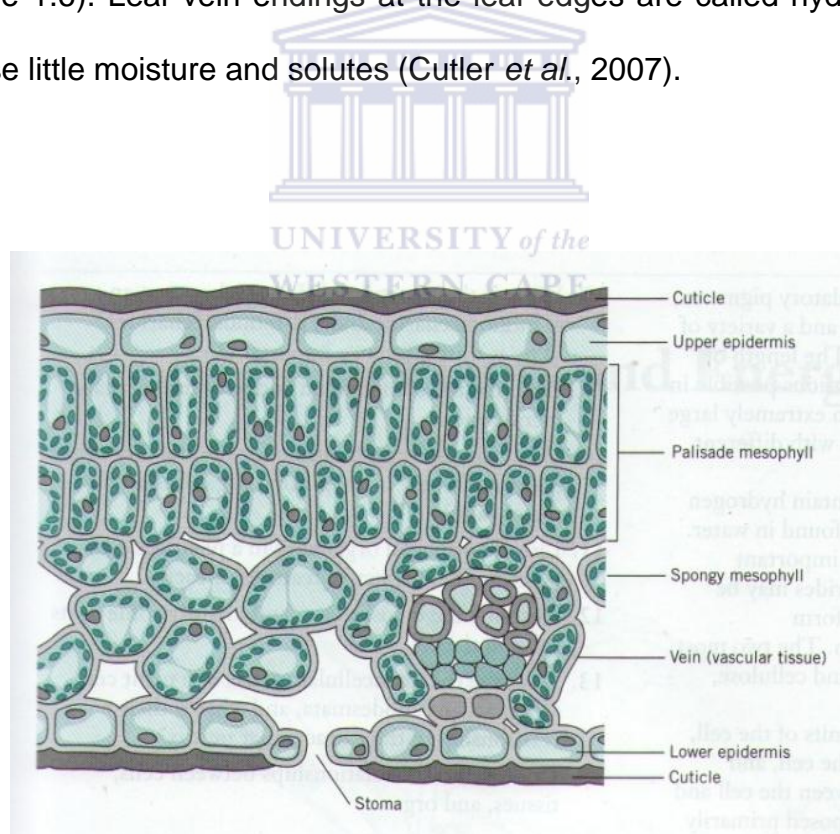
Figure 1.6. Diagrammatic cross-section illustrating the principal features of a typical dicotyledonous leaf. (Adapted from Hopkins *et al.*, 2004)

The next sections discuss, in greater detail, the establishment of endophytic bacterial communities. The different plant-microbe interactions are described, as well as their ecological and biotechnological importance.

1.1.3. Recruitment and adaptation of endophytic bacterial communities

1.1.3.1. Recruitment of endophytic bacteria

Endophytic bacterial communities differ in structure to communities of surrounding soils. This is due to the selectivity of the lateral transmission process and the presence of vertically transmitted bacteria. Recruitment of endophytic communities takes place mostly through the roots from the rhizosphere (Lugtenberg and Kamilova, 2009). All bacteria found in the vicinity of the root surface have the opportunity to invade plant tissues. Bacteria form microcolonies at the entry "hotspots" such as the zones of elongation and differentiation, surface wounds and "cracks" at the points of emergence of lateral roots (Walker *et al.*, 2003). At these points, bacteria can enter the plant tissues through the ridges between the epidermal cells. Figure 1.7 illustrates different modes of infection by endophytic bacteria.

"Passenger/transient" endophytes do not invade the plant beyond the root pericycle (Figure 1.7). These are retained in the endosphere for short periods as they are unable to colonise plant tissues and establish communities in this environment (Hardoim *et al.*, 2008). Opportunistic endophytes are retained for longer periods within the root tissues due to minimal selective forces, and they can even multiply

(Hardoim *et al*, 2008). Only competent endophytes (also called "true" endophytes) can establish communities in the endosphere even when selective pressure is high (Hardoim *et al.*, 2008). Figure 1.7 shows that once inside the root cortical tissues, true endophytes are able to move away from the zone of entry and invade other tissues, some even entering the vascular system. True endophytes include members of α-*Proteobacteria*, β-*Proteobacteria*, γ-*Proteobacteria*, *Firmicutes*, *Bacteroidetes* and *Actinobacteria* (Rosenblueth and Martínez-Romero, 2006).
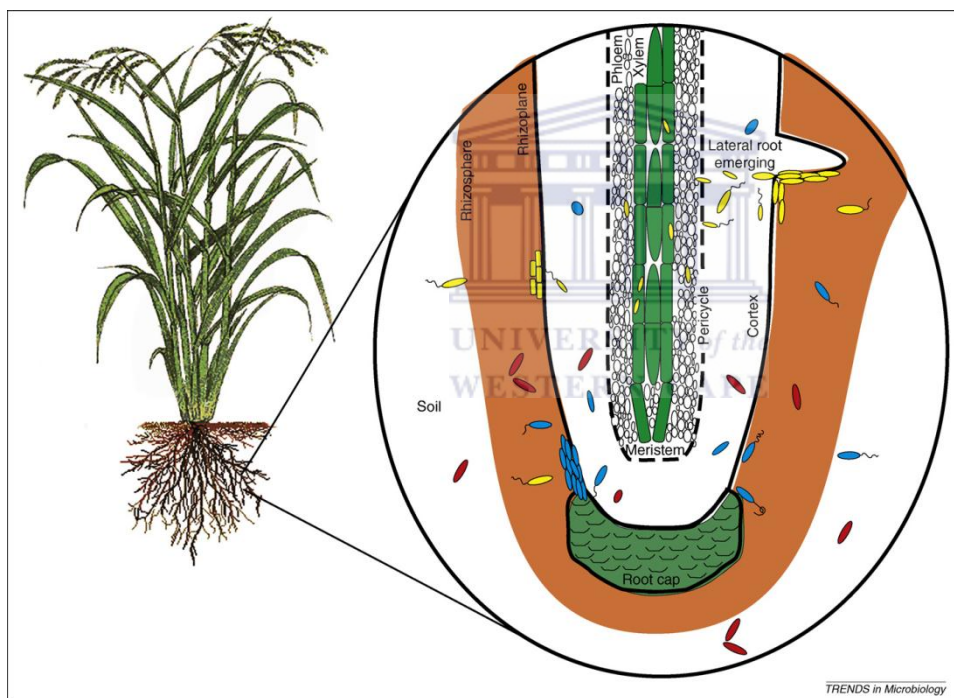
Figure 1.7. Types of endophytes and their root colonization process. Stochastic events and deterministic bacterial factors drive colonization of the endosphere, in which a series of events, including microcolony formation at the root surface, are thought to take place. Bacteria entering plant tissues could be competent (yellow) opportunistic (blue cells) or passenger endophytes (red cells). (Adapted from Hardoim *et al.*, 2008)

1.1.3.2. Adaptation strategies for colonisation of the endosphere

Microbes invading plant tissues face several adaptive challenges for them to colonise and establish communities in this environment. The plant structures are obstructive and can prevent microorganisms from accessing their niche of optimal growth. Plant cells are connected by cellulosic cell walls, and plant tissues such as the xylem are reinforced with rigid lignin (Hopkins *et al.*, 2004). Microorganisms also have to overcome the plant's defense system against foreign bodies. The plant's defence system recognises foreign cells in its tissues, and this leads to a response that involves increased production of antimicrobial phytohormones such as ethylene, salisylic acid and jasmonic acid (Chisholm *et al.*, 2006; Hardoim *et al.*, 2008). Moreover, endophytic bacteria need to have access to essential nutrients to grow in such environments, and this may lead to competition for space and nutrients between endophytic organisms. Therefore, the list below highlights key adaptive attributes of true endophytes (as reviewed by Hardoim *et al.*, 2008):

1. <u>Chemotaxis</u>: The chemical recognition patterns and chemotactic responses are controlled at the genetic level in endophytic bacteria. These enable recognition of plant exudates components by the bacteria and movement towards the root surface where recruitment into the plant tissues takes place (Reinhold-Hurek and Hurek, 2011).

2. <u>Production of adhesion molecules</u>: Infection of the plant by endophytic bacteria takes place at the root-soil interface. As such, these bacteria produce adhesive polysaccharides in order to attach to the root surface and form microcolonies.

3. Mobility: True endophytes possess mobility structures such as Type IV pili and flagella (Wieland *et al.*, 2001), and these structures enable bacteria to move from the area of invasion towards their niche.

4. Production of plant polymer-degrading enzymes: Endophytes producing enzymes such as endoglucanases, polygalacturonidases, cellulases and xylanases are able to digest obstructive plant structures and move through the plant tissues (Cho *et al.*, 2007).

5. Suppression of the plant's defense system: Endophytic bacteria have the ability to modulate production of antimicrobial phytohormones produced by the plant, thus neutralising the potency of the immune response. For examples, endophytic bacteria can interrupt production of ethylene by releasing 1-aminocyclopropane-1-carboxylate synthase (ACC) deaminase, an enzyme that degrades ACC, the immediate precursor of ethylene.

6. Lack of plant immune response elicitors: Surface molecules and metabolites (such as proteins released by the Type III and Type IV protein secretion system) common to most pathogens are recognised by the plant, and these trigger an immune response. Some endophytes lack these elicitors, and are therefore able to avoid detection by the plant's defense system.

7. Competition: Competent endophytes have to be effective competitors against other organisms living in the endosphere. For example, endophytes that produce siderophores are able to sequester and utilise iron, which is a limiting micronutrient in the endosphere. Siderophore producers can colonise the iron-deficient endosphere more effectively than non-siderophore producers.

8. Symbiosis potential: It is hypothesised that some host plants have evolved to preferentially select for microorganisms that have beneficial properties for

plant growth and health. When these microorganisms invade plant tissues, the immune response is minimal, thus giving the microorganisms a competitive advantage over other plant tissue inhabiting microorganisms (Vargas *et al.*, 2012).
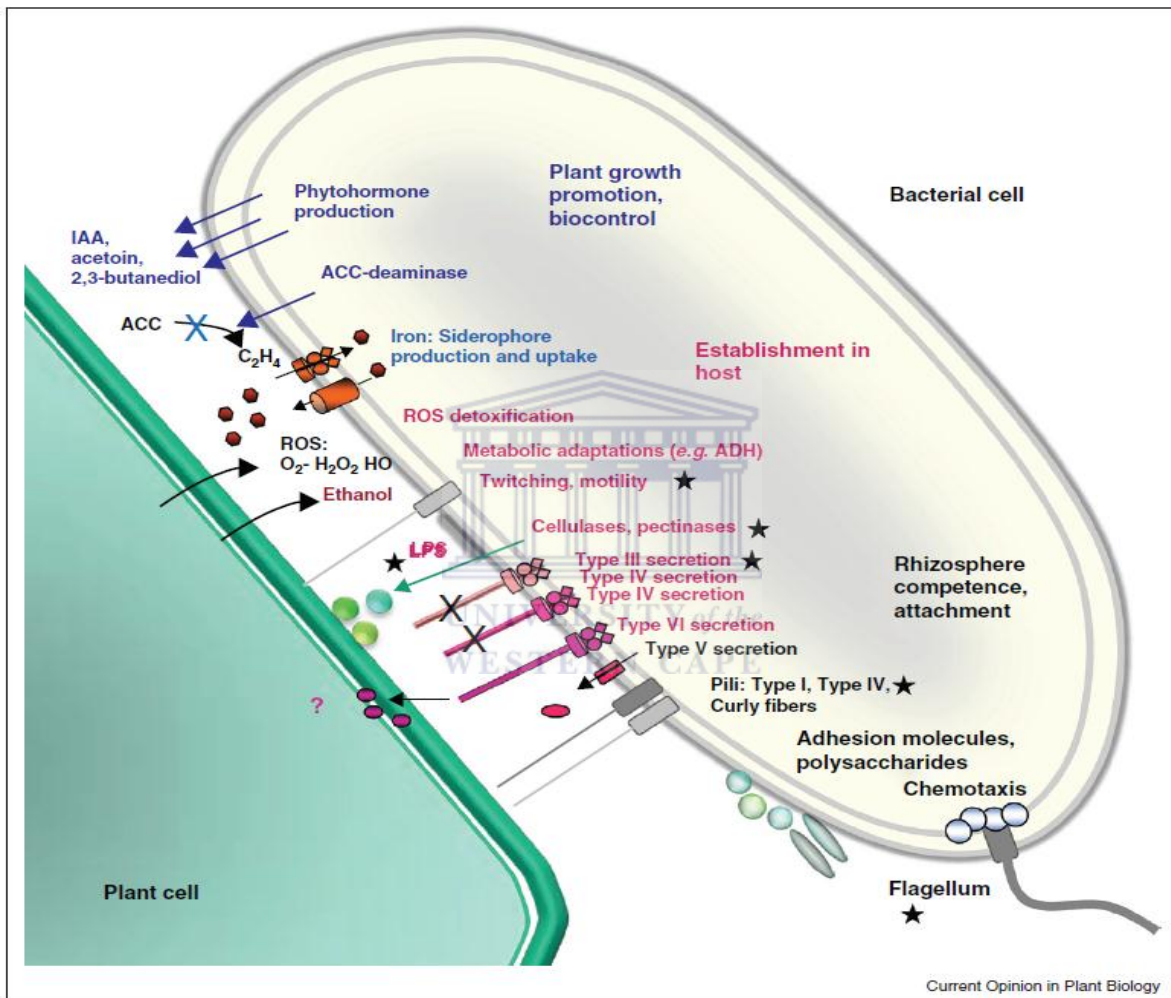
Figure 1.8. Examples for bacterial characteristics putatively involved in endophyte–plant interaction, as shown by experimental (mostly mutational) studies (labeled with star), or suspected inferred from literature context or genome comparisons (Adapted from Reinhold-Hurek and Hurek, 2011).

True endophytes undergo various metabolic processes during their progress from the external environment to the inner plant tissues where they establish communities. This indicates that recruitment of endophytic communities is a highly selective process that requires activation and deactivation of several genetically-controlled pathways as summarised in Figure 1.8.

1.1.4. Plant-microbe interactions in the endosphere

The close association between the host plant and the microorganisms living in the plant-created microenvironments results in the establishment of diverse and complex interactions (Figure 1.9), which are important to microbial and plant life. As previously mentioned, microorganisms benefit from these associations as the plant provides them with habitat for colonisation and nutrients. The broad range of microbial metabolic capabilities can be beneficial or detrimental to the plant.

1.1.4.1. Beneficial interactions

PGPMs enhance the fitness of the host plant by increasing availability of essential nutrients to the plant (Vessey, 2003), production of plant-growth inducing hormones (Dodd *et al.*, 2010), enhancing the plant's defence system (Heydari and Pessarakli, 2010) and/or remediating soils on which the plants grow (Smith *et al.*, 2010). This mutually beneficial interaction between plants and microorganisms is also known as symbiosis (Newton *et al.*, 2010). The following are some of the well-understood interactions involving plant growth promoting endophytic bacteria (PGPEBs).
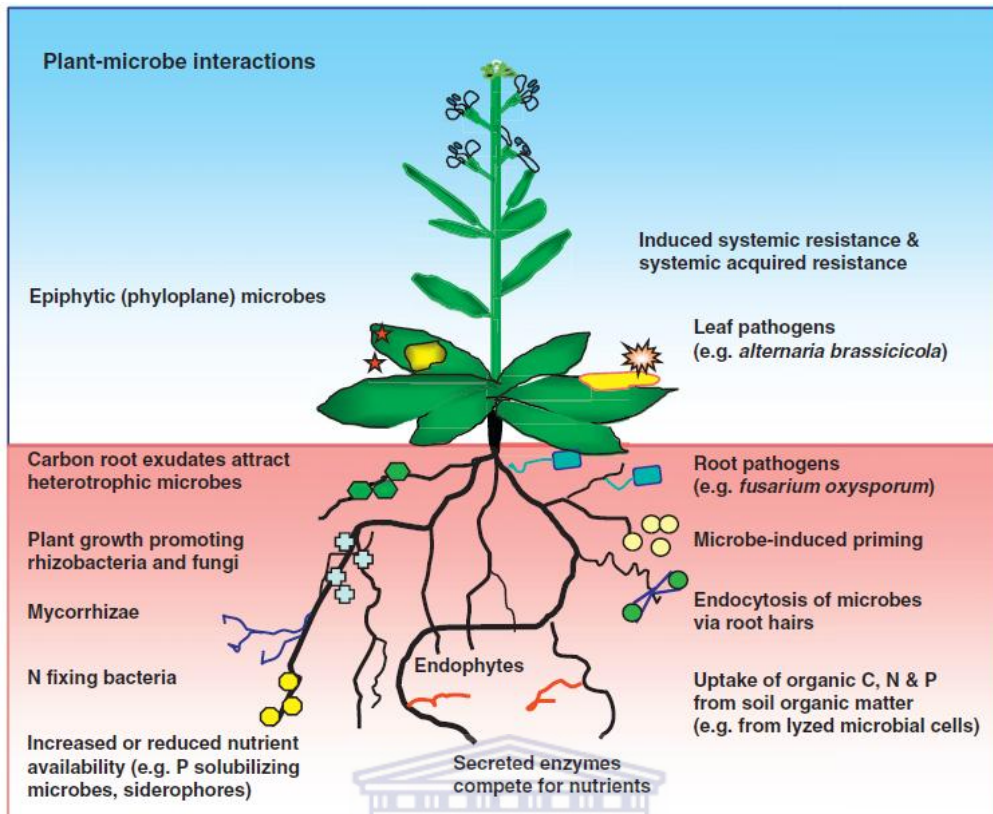
Figure 1.9. General overview of interactions between plants and microbes (Adapted from Schenk *et al*., 2012).

a. Nitrogen-fixing endophytic bacteria

Nitrogen is a limiting nutrient for plant growth. Sources of nitrogen in the soil are mostly in the form of atmospheric dinitrogen and insoluble soil nitrogen compounds, and both are not readily available for assimilation by plants. Diazotrophic bacteria are able to convert molecular nitrogen to bioavailable ammonium through a nitrogenase-mediated reaction. Production of the nitrogenase enzyme is encoded by the bacterial *nif* genes (Kraizer *et al*., 2011). Diazotrophic endophytes are considered beneficial to the plant only when the fixed nitrogen is transferred to the host plant as not all of them have been shown to do so (James, 2000). The most studied and

important nitrogen-fixing endophytes are the nodule-forming Rhizobia associated with legume plants such as soybean and groundnut (Kraizer *et al.*, 2011). Several members have been isolated and characterised (Figure 1.10).
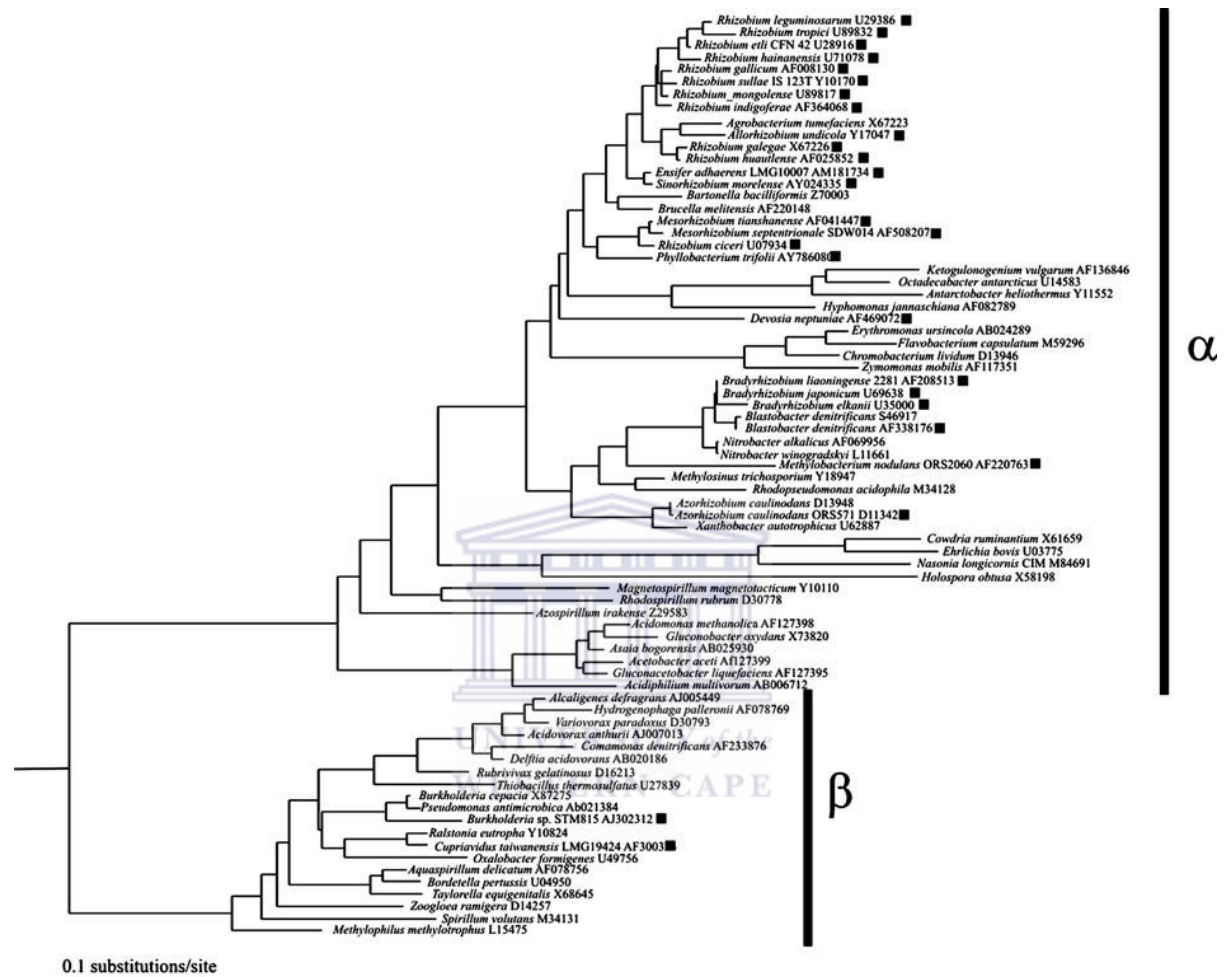
Figure 1.10: Phylogeny of Rhizobia. A maximum likelihood tree based on *rrs* genes from 75 taxa from alpha- and betasubdivisions of Proteobacteria. Representatives of species capable of forming nodules are marked with a black box (Adapted from Franche *et al.* 2009).

Another group of nodule-forming - yet less understood - nitrogen-fixing endophytic bacteria are the slow-growing *Frankia* species (Franche *et al.*, 2009). The family

Nostocales also includes some endophytic nitrogen-fixing cyanobacterial species. Under nitrogen-deprivation conditions, these filamentous cyanobacteria express genes such as *het*, *ntc* and *pat* that are responsible for heterocyst formation. The heterocysts are a site of nitrogen fixation, and fixed nitrogen can be transported to other parts of the plant (Franche *et al*., 2009).

Endophytic nitrogen-fixing bacteria that do not form nodules or heterocysts have been isolated in non-legume plants (Reis *et al*., 2011). *Azoarcus* sp. BH72 has been isolated from monocotyledonous plants such as rice plants (Hurek *et al*., 1994; 1998; Reinhold-Hurek and Hurek, 1998a; Reinhold-Hurek and Hurek, 1998b) and kallar grass (Malik *et al*., 1997). Expression of *nif* genes in rice roots by *Azoarcus* sp. BH72 was confirmed, and these bacteria were also shown to increase the level of biologically fixed nitrogen in plant tissues (Hurek *et al*., 2002). Complete genome sequences of diazotrophic grass endophytes *Azoarcus* BH72 and *Gluconacetobacter diazotrophicus* Pal5 are currently available (Krause *et al*., 2007; Bertalan *et al*., 2009), thus making these microorganisms model organisms for future studies of plant-associated diazotrophs.

b. <u>Phosphorus-solubilising bacteria</u>

Phosphorus (P), as nitrogen, is an essential macronutrient required by plant cells for the production of various structural and functional metabolic products such as nucleic acids (Madigan *et al*., 2009). However, most of the soil phosphate is inaccessible to plants. Plant cells are only able to take up soluble phosphate of low

molecular weight such as ionic P (Pi, HPO^{2-} and H_2PO^-) or low molecular organic phosphate (Rodriguez and Fraga, 1999).

Microbial activities play a key role in liberating phosphorus from organic and inorganic compounds. This is achieved by solubilisation of insoluble inorganic P compounds and mineralization of organic compounds by a group of microorganisms referred to as phosphate solubilising microorganisms (PSMs) (Podile and Kishore, 2006; Khan *et al.*, 2009). In the solubilisation of inorganic phosphate compounds, PSMs release protons and organic acids which lower the environmental pH, and thus increase the solubility of phosphate compounds which then become available to plants (Khan *et al.*, 2009). PSMs can also hydrolyse phosphorus-containing molecules such as calcium phosphate by releasing organic acids or hydrolytic alkaline phosphatases (Khan *et al.*, 2009).

Some of the well-known microorganisms with P-solubilising abilities include strains from the bacterial genera *Pseudomonas*, *Bacillus*, *Rhizobium* and *Enterobacter,* and fungal strains of *Penicillium*, *Aspergillus* and mycorrhizal fungi (Khan *et al.*, 2009). Phosphate solubilising bacteria are often abundant in plants growing in phosphate-deprived soils. This was shown by Kuklinsky-Sobral *et al.* (2004), whereby a total of 373 endophytic bacterial isolates – mostly from the families Pseudomonaceae, Burkholderiaceae and Enterobacteriaceae – from soybean cultivars were able to solubilise mineral phosphate. These phosphate-solubilising endophytes were most predominant in the earlier stages of the plant's development, thus suggesting that plants have greater need for this limiting macronutrient for metabolic and structural purposes in their vegetative stage.

c. Siderophore-production

Microorganisms that produce siderophores play an important role in the plant's iron-acquisition and defense against pathogens. Siderophores are soluble, low-molecular weight compounds that are able to chelate iron in the environment (Saha *et al.*, 2012). Iron is abundant in the soil; however, it is mostly oxidised to the biologically unavailable ferric compounds (Saha *et al.*, 2012). Bacteria that produce siderophores, including *Pseudomonas*, *Bacillus*, *Micrococcus* and *Variovorax* strains (Sun *et al.*, 2011; Rashid *et al.*, 2012), release these molecules into the environment where they chelate iron and make it available to the bacteria and the host plant (Miethke and Marahiel, 2007; Saha *et al.*, 2012). Siderophores such as those produced by the *Pseudomonas* sp. strain MW2.6 have also been shown to have antimicrobial properties against plant pathogens including *Alternaria*, *Fusarium oxysporum*, *Pyricularia oryzae* and *Sclerotium* (Chaiharn *et al.*, 2009).
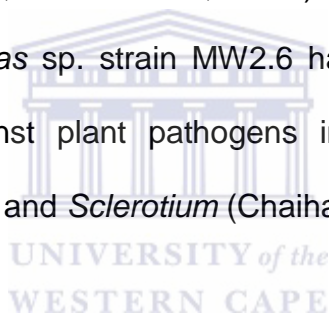
d. Production of plant-growth inducing hormones

The main groups of plant-growth promoting phytohormones are auxins, gibberellins (GAs), abiscisic acid (ABA), cytokinins and ethylene, and these are produced by the plant through complex, genetically-controlled pathways throughout its development (Vandenberghe *et al.*, 2014). Auxins (e.g. indole acetic acid [IAA]), cytokinins and GAs are implicated in cell division and growth, seed germination, leaf growth and tissue differentiation whilst ABA promotes seed maturation. Ethylene is produced at the mature stage of the plant's development to facilitate fruit ripening, abscission and senescence (Vandenberghe *et al.*, 2014). Stress-induced interruption of

phytohormone production pathways in plants thus leads to stunted plant growth and poor tissue development (Vanneste and Friml, 2009).

Production of growth-inducing phytohormones has also been observed in plant-associated bacteria; therefore these microorganisms can directly promote plant growth. Previously isolated and characterised bacteria that have been reported to produce phytohormones include, among others, strains of *Achromobacter xylosoxidans*, *Bacillus subtilis*, *Athrobacter koreensis*, *Microbacterium testaceum* (Forchetti *et al*., 2007; Jha and Kumar, 2009; Gagne-Bourgue *et al*., 2011; Malfanova *et al*., 2011; Piccoli *et al*., 2011)
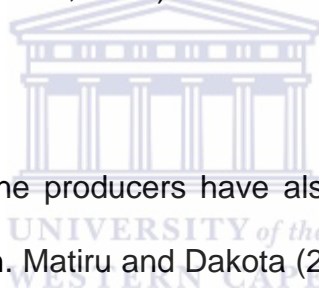
Isolated potential phytohormone producers have also been inoculated in plants to test their effect on plant growth. Matiru and Dakota (2004) showed that sorghum and millet roots were easily infected by phytohormone producing rhizobacteria, *Azorhizobium caulinodans* ORS571, *Rhizobium* GRH2 and *Bradyrhizobium japonicum* TAL110, which have also been shown to increase plant growth. Furthermore, ABA-, IAA- and GA- producing *Azospirillum lipoferum* was shown to improve growth of maize plants with inhibited plant-meadiated ABA and GA synthesis pathways under drought conditions, thus indicating that production of phytohormones by endophytic bacteria can alleviate abiotic stress in host plants (Cohen *et al*., 2009).

e. Endophytes with biocontrol potential

Plant parasites and pathogens deplete nutrients, and can also cause diseases that hinder the plant's development. Antagonistic endophytes release antibiotics and other chemicals (e.g. siderophores) that kill pathogens or slow their colonisation in the plant's tissues, thus enhancing the plant's resistance to disease. For example, inoculation of *Pseudomonas chlororaphis*, *Lysobacter gummosus*, *Paenibacillus polymyxa* and *Seratia plymuthica* in Styrian oil pumpkins showed significant antagonism against the fungal pathogen *Didymella bryoniae* (Fürnkranz *et al.*, 2011). In greenhouse in vivo trials carried out by Aravind *et al.* (2009), endophytic bacterial isolates, *Pseudomonas aeruginosa*, *P. putida* and *Bacillus megaterium* significantly suppressed growth and development of *Phytophthora capsici* that causes foot rot in black pepper plants.

1.1.4.2. Non-beneficial interactions

In these interactions, the host plant does not benefit from its association with microorganisms. One type of this association is called commensalism, whereby microorganisms infect plants without causing apparent symptoms of disease or providing benefits to the plant (Newton *et al.*, 2010).

Parasitic and pathogenic microorganisms are harmful to plant life because they cause plant disease and death. Parasitism occurs when microorganisms utilise the plant's resources such as water and nutrients at the expense of the plant's health, growth and development. Depletion of the plant's resources reduces its fitness and

increases its susceptibility to diseases (Newton *et al.*, 2010). A classic example of parasitism occurs when the fungal species, *Golovinimyces orontii,* infects the leaves of *Arabidopsis thaliana.* Fungal cells develop a structure called a haustorium, which is a specialised long cell that penetrates plant cells and feed from the cytoplasm (Micali *et al.*, 2011).

Pathogenesis takes place when microorganisms feed on plant tissues, often leading to necrosis, which is unprogrammed death of plant cells/tissue (Newton *et al.*, 2010). Biotrophic pathogens such as the endophytic *Pseudomonas syringae* that infects *Arabidopsis thaliana*, feed on living plant tissue (Butt *et al.*, 1998), whereas necrotrophic pathogens first induce death of plant tissues and then feed on dead matter (Glazenbrook, 2005). *Botrytis ceneria* is an example of a necrotrophic fungal pathogen, which was shown to release phytotoxins to induce death of host plant tissue shortly after infection (Colmenares *et al.*, 2002).

Plant-microbe interactions are considered to be dynamic as they exist as a continuum between the two extremes of mutualism and pathogenesis (Figure 1.11) (Newton *et al.*, 2010). A plant associated microorganisms can enter different states of trophism in its life cycle, and changes between trophic spaces are often in response to environmental, host development and microbe-specific triggers (Newton *et al.*, 2010). Such a change between trophism states is exhibited by the obligate endophyte *Herbaspirillum rubrisubalbicans* that infects agriculturally important Poaceae grasses including sugarcane, sorghum, millets, wheat, rice and maize. This endophyte was also shown to cause mottled stripe disease in some varieties of

sorghum and sugarcane (James *et al.*, 1997). However, the pathogenic state can alter to a mutualistic state by altering the type III secretion system in *H. rubrisulbalbicans*, and it can exhibit plant-growth promoting properties such as nitrogen fixation in plants (Schmidt *et al.*, 2012). Endophytic bacterium, *Helicobacter pylori,* alternates between commensal and pathogenic trophic states depending on the host plant species it is associated with (Hirsh, 2004).
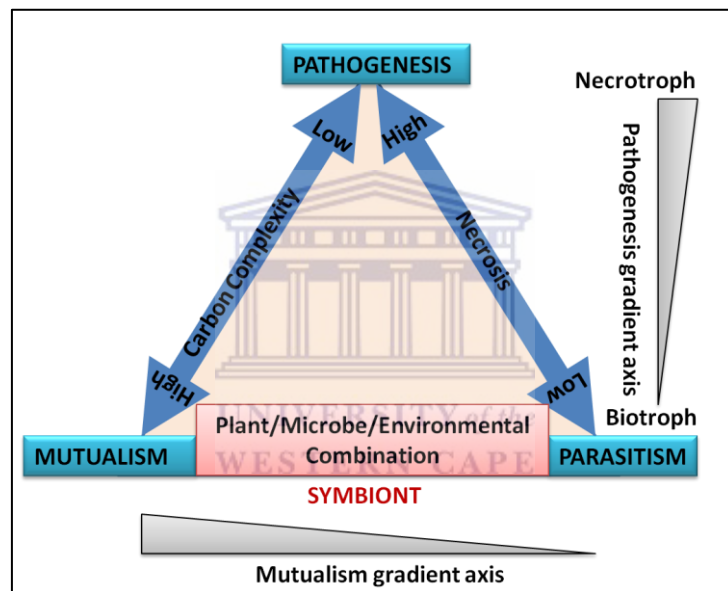
Figure 1.11. Trophic spaces occupied by plant-associated microorganisms. Plant, microbe and environmental cues determine the trophic space occupied by the microorganism at any phase of its lifecycle. The vertical pathogenesis axis and the horizontal mutualism axis grade the cost and benefit of the microbial association to the plant, respectively (Adapted from Newton *et al.*, 2010).

1.1.5. Importance of plant-microbe interactions

Plant-microbe interactions play a crucial role in soil ecosystems as well as plant health and growth. Understanding these interactions is thus important in agronomical management of plant disease and soil quality, and they can also be exploited for biotechnological developments and industrial applications in various fields.

1.1.5.1. Importance to plant life

PGPMs have a positive impact on plant health and growth whereas pathogens and parasites decrease the plant's fitness and cause disease (Newton *et al.*, 2010). Beneficial properties of PGPMs are currently exploited in agricultural applications for the production of biofertillizers and biocontrol agents (Andrews *et al.*, 2010). Biofertilizers are living beneficial microorganisms that colonise the rhizosphere or the endosphere, and promote plant growth (Vessey, 2003). For example, *Achromobacter xyloxidans* (a wheat endophyte) has the potential for use as a biofertilizer due its plant growth promoting properties that include nitrogen fixation, phosphate solubilisation and the production of IAA (Jha and Kumar, 2009).

Biocontrol agents are living microorganisms that have deleterious effects on plant pathogens and pests (Andrews *et al.*, 2010). They can be applied to either soils and/or plants in order to improve plant health. Microorganisms with antibiosis properties, mycoparasites and effective competitors in the biosphere have the potential as biocontrol and/or biopesticide inocula (Heydari and Pessarakli, 2010). *Paenibacillus polymyxa* GS01, *Bacillus* sp. GS07 and *Pseudomonas poae* JA01

isolated from ginseng roots showed significant activity against phytopathogens including *Rhizoctonia solani*, *Phythium ultimunn*, *Fusarium oxysporum* and *Phytophthora capsici*, and could therefore be considered for use as biocontrol agents in management of crop diseases caused by these pathogens (Cho *et al.*, 2007).

Biofertilizers and biocontrol agents are nowadays available commercially. They provide environmentally-friendly alternatives to the non-biodegradable fertilizers and pest-control chemicals that lead to soil quality deterioration when used for prolonged periods (Kennedy and Smith, 1995; Lugtenberg and Kamilova, 2009). Notably, Mycoroot is a South African company that provides a selection of arbuscular mycorrhizal fungi biofertilizer inocula for increased uptake of nutrients and water in plants and improved soil structure and aeration (Mycoroot™). Also, Actinovate ® AG is a broad range commercial fungicide that contains *Streptomyces tydicus* for control of foliar fungal phytopathogens (Natural Industries Inc.). Table 1.1 lists other examples of endophytic bacteria that can be considered for production of biofertilisers and biocontrol agents.

Table 1.1. Endophytic bacteria with biofertiliser and/or biocontrol potential.

Strain	Source	Property	Benefits to plants	Reference
Pseudomonas sp. AKM P6	pigeon pea rhizosphere	Biofertilizer	Phytohormone (IAA, gibberellins) production. Increased growth and thermotolerance of sorghum plants.	Ali *et al.*, 2009
Pantoea agglomerans YS19	Rice	Biofertilizer	Phytohormone (IAA, abscisic acid, gibberellins and cytokinins) production, nitrogen fixation, promotion of effective allocation of photosynthetic products within infected rice plant tissues.	Feng *et al.*, 2006
Bacillus sp. SLS18	Pokeberry (*Phytolaca acinosa*)	Biofertilizer Bioremediation agent	Phytohormone (IAA) production, siderophore production Increases plant (sorghum) tolerance to heavy metals (Mn/Cd)	Luo *et al.*, 2011
Gluconacetobacter diazotrophicus Pal5	sugarcane	Biofertilizer	Nitrogen fixation, phytohormone production, phosphate and zinc solubilisation	Sevilla *et al.*, 2001
		biocontrol agent	Antibacterial activity against sugarcane pathogen *Xanthomonas albilineans*, and antifungal activity against corn pathogens *Fusarium* sp. and *Helminthosporium carbonum*	Blanco *et al.*, 2005; Mehnaz and Lazarovits 2006.
Pantoea ananatis 125NP12	tomato	biocontrol agent	IAA production	Enya *et al.*, 2007
		Biofertilizer	Antifungal activity against *Botrytis cinerea*, *Fulvia fulva*, and *Alternaria solani*	

1.1.5.2. Importance to microbial life

The autotrophic host plant produces energy-rich carbon compounds by photosynthesis and release nutrient-rich metabolites into the different plant-created microenvironments (Cutler *et al.*, 2007). These plant-produced compounds serve as food for plant-associated microorganisms, and are thus necessary for the establishment of microbial communities. In addition to the relationships established between the host plant and microorganisms, other important interactions in these microenvironments include those between different microbial species themselves (Rosenblueth and Martínez-Romero, 2006). Microorganisms that inhabit the same microbial niche within a plant-created microenvironment could compete for nutrients and space. This was observed in a study where a fast growing *Pantoea* sp. minimised colonisation of *Ochrobactum* sp. in the rice plant tissues (Verma *et al.*, 2004).
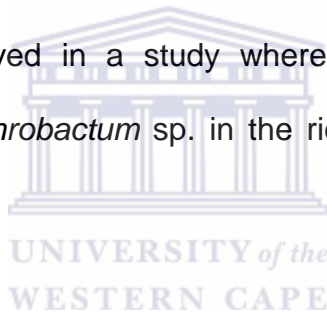
Mutualistic relationships have also been observed between plant-associated microorganisms. A case in point is the observed relationship between an endophytic Bacillus species and arbuscular mycorrhizal fungi in Brazilian arnica roots (*Solidago chilensis*) (Silvani *et al.*, 2008). The hyphae of the mycorrhizal fungi grow and branch between plant cells where they are able to absorb plant metabolites. These hyphae extend beyond the plant root tissue into the mycorrhizosphere where they sequester nutrients such as phosphorus and make them available to the plant, promote soil aggregation for increased water retention capacity and facilitate remediation of metal-polluted soils (Hodge and Andrews, 2004; Rillig, 2004; Bedini *et al.*, 2009). Inside the plant roots, the hyphae provide a growth substrate for *Bacillus* sp., which

were also able to access nutrients from the fungus (Figure 1.12) (Silvani *et al.*, 2008). The endophytic *Bacillus* species is implicated in production of plant-growth promoting hormones, which increase root biomass; thus indirectly expanding the habitat for fungal colonisation (Silvani *et al.*, 2008). This complex, interdependent relationship between plants and microorganisms also highlights the ecological importance of plant-associated microorganisms in the broader ecosystem, which should be considered in the management of plant health and growth.
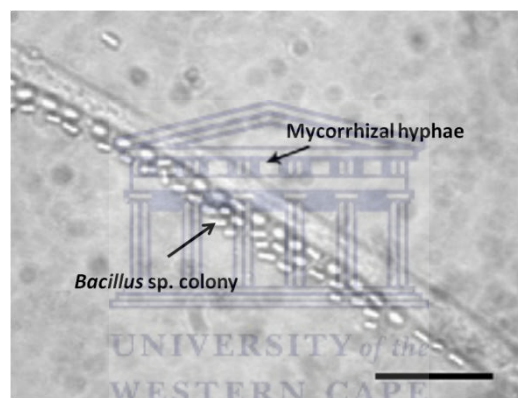
Figure 1.12. *Bacillus sp.* growing along the hyphae of an arbuscular mycorhizal fungus found within the root of *Solidago chilensis* (Brazilian arnica) viewed using a binocular microscope. Bar=200μm. (Adapted from Silvani *et al.*, 2008).

1.1.5.3. Biotechnological importance of plant-associated microorganisms

Metabolic capabilities of the plant-associated microorganisms have great potential for biotechnological exploitation (Ryan *et al.*, 2007; Compant *et al.*, 2010; Qin *et al.*, 2011). Plant-associated fungi and bacteria produce natural products used in different

industrial applications. Table 1.2 lists some examples of natural products derived from endophytic bacteria.

Table 1.2. Natural bioproducts derived from various endophytic bacteria. (Adapted from Ryan *et al.*, 2007)

Organism	Plant association	Active agent	Activity	Reference
Taxomyces andreanae	*Taxus brevifolia*	Taxol	Anticancer	Strobel *et al.* (1993)
Pseudomonas viridiflava	Grass	Ecomycins B and C	Antimicrobial	Miller *et al.* (1998)
Streptomyces griseus	*Kandelia candel*	p-Aminoacetophenonic acids	Antimicrobial	Guan *et al.* (2005)
Streptomyces NRRL 30562	*Kennedia nigriscans*	Munumbicins Munumbicin D	Antibiotic Antimalarial	Castillo *et al.* (2002)
Streptomyces NRRL 30566	*Grevillea pteridifolia*	Kakadumycins	Antibiotic	Castillo *et al.* (2003)
Serratia marcescens	*Rhyncholacis penicillata*	Oocydin A	Antifungal	Strobel *et al.* (2004)
Paenibacillus polymyxa	Wheat, Lodge pine, Green beans *Arabidopsis thaliana* Canola	Fusaricidin A–D	Antifungal	Beck *et al.* (2003) Li *et al.* (2007) Beatty & Jensen (2002)
Cytonaema sp.	*Quercus* sp. 103	Cytonic acids A and D	Antiviral	Guo *et al.* (2000)
Streptomyces sp.	*Monstera* sp.	Coronamycin	Antimalarial antifungal	Ezra *et al.* (2004)

Natural products that have antipest or antimicrobial properties can be used in the medical field for the production of antibiotics or in industry for the production of disinfectants and agricultural pesticides (Gunatilaka, 2006). Some natural products are antioxidants and some have anti-cancer properties. Other microorganisms or microbial products have been used for pollution control and phytoremediation processes (Ryan *et al.*, 2007). Figure 1.13 shows the vast applications of plant-associated microorganisms and their products.
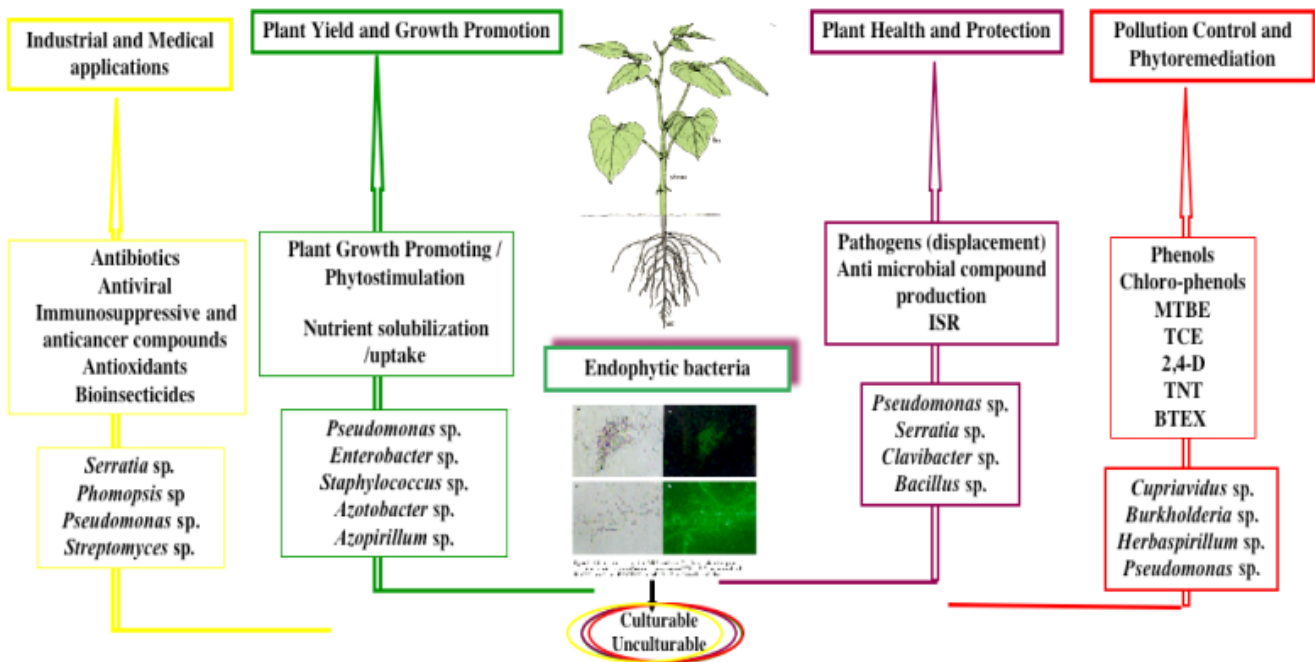
Figure 1.13. Schematic representation of various plant-endophyte interactions and their applications (Ryan *et al*., 2007).

The next section reviews methods used in the study of endophytic bacterial communities, and discusses some of the major advances made in this field.

1.2. Studying endophytic bacteria: understanding their community structure, biology and interactions in the endosphere

The first microorganisms observed within plant tissues were fungi at the end of the 19th century, but "endophytes" were only first described in Darnel in 1904 (Kusari *et al*., 2012). Since then, endophytic microorganisms were isolated from various plants, with research geared towards understanding the establishment of these microbial

communities and their ecological importance. Currently, there is a broad selection of culture-dependent and culture-independent tools available for the study of endophytic microbial communities.

1.2.1. Culture-dependent approaches in the study of endophytic communities

Conventional microbiological culturing approaches have been used to isolate endophytic bacteria from surface-sterilised plant tissues. Non-selective media can be used to enumerate the entire culturable community, whilst selective media are used for targeted isolation of specific groups of endophytes. For example, nitrogen-free media such as LGI, NFB and JNFb- were developed for specific isolation of diazotrophic bacteria (Kirchhof *et al.*, 1997), whereas phosphate-solubilising bacteria can be cultured on media such as PVK and NBRIP (Nautiyal, 1998).

Once isolated, these strains can be identified by sequencing and/or subjected to various phenotypic and biochemical characterisations to test plant growth promoting capabilities such as the production of essential biomolecules with potential biotechnological applications (e.g. antibiotics, polymer-degrading enzymes, plant-growth hormone) (Enya *et al.*, 2007; Ulrich *et al.*, 2008; Silvani *et al.*, 2008; Magnani *et al.*, 2010). These culture-based studies are important as they can give an indication of possible plant growth promoting activities taking place in the endosphere, and can lead to the discovery of novel bacterial species and/or relevant biomolecules. Isolated strains can also be manipulated by genetic engineering

processes to increase production of specific metabolites, or to increase targeted activity. However, a major limitation of culture-based studies is that less than 1% of microorganisms in the biosphere can be cultured on available media (Morgan and Whipps, 2001); therefore molecular tools are necessary to study the diversity and functions of endophytic microbial communities (Hurek *et al.* 2002; Sessitsch *et al.*, 2012).

1.2.2. Culture-independent approaches in endophytic bacterial community studies

Developments in molecular biology have made it possible to study phylogenetic assemblages of bacterial communities directly from their natural environment as well as to analyze various functional activities in the whole community (Sessitsch *et al.*, 2012).
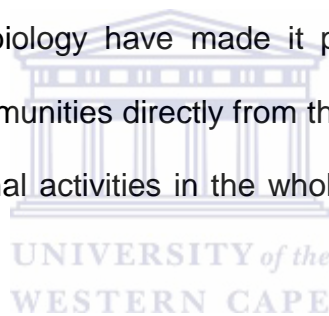
1.2.2.1. Bioassays

Certain metabolites and surface structural components are unique to specific microbial groups. A survey of these biomolecules from environmental samples can assist in the estimation of parameters such as population abundance and composition. The fatty acid methyl esters (FAME) and phospholipid fatty acid (PLFA) assays have been used to determine the microbial community structures from complex environments (Parekh and Bardgett, 2002). FAME, for instance, was used in the identification of 140 endophytic bacterial species from coffee plant fruit (*Coffea canephora*) (Miguel *et al.*, 2013). PLFA has been used to study the effect of

Huanglongbing (HLB) disease on bacterial and fungal community structures in HLB-affected leaves of red pomelo, a Chinese citrus plant (Zheng *et al.*, 2012).

1.2.2.2. Metagenomic community analyses

Microbial community profiles can be estimated using genomic material of the entire microbial population extracted directly from their natural habitat. Metagenomic DNA extracted from complex environments including endophytic environments has been used to determine more accurately the composition and diversity of microorganisms living in those environments using metagenomic tools (Marschner *et al.*, 2005), some of which are described hereafter.
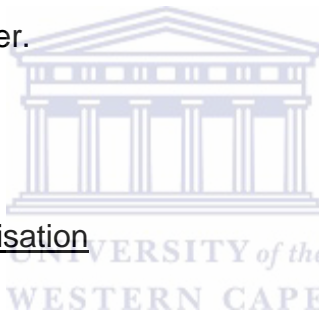
a) DNA:DNA hybridisation

In this technique, radio- or isotope-labelled DNA probes of known sequences – often complimentary to specific functional/structural sequences or phylogenetic sequences – are exposed to metagenomic DNA and the level of hybridisation is measured and analysed (Pareck and Bardgett, 2002). The information obtained can be used to determine the metabolic potential and/or phylogenetic structure of the bacterial community found in that environment. Fluorescent *in situ* hybridisation (FISH) uses mDNA or mRNA probes that can penetrate bacterial cells whilst retaining their morphological structure, and can be used to identify bacteria in their environment (Pareck and Bardgett, 2002). FISH was used to confirm the establishment of a tripartite association between pine trees, endophytic fungi and endophytic bacteria (Figure 1.14).
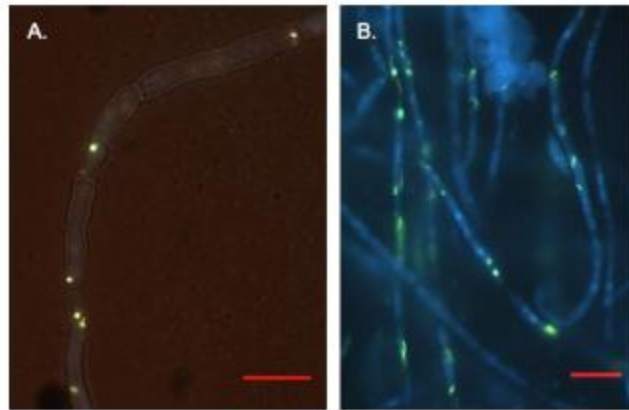
Figure 1.14. Fluorescent *in situ* hybridization (FISH) microscopy showing hyphae of two isolates of endophytic fungi harboring endohyphal bacteria. Panel A (isolate 9084b; Dothideomycetes) shows the TAMRA fluorophore at the site of internal structure in hyphal cells. Panel B (isolate 9143; Sordariomycetes) shows the TAMRA fluorophore with the DAPI counterstain (blue), highlighting the fungal nuclear and mitochondrial DNA in addition to bacteria (yellow/green). Scale bar, 10 μm (A) or 25 μm (B). (Adapted from Hoffman and Arnold, 2010).

b) Community profiling techniques

The distribution of structural, genomic and functional biomolecules in the environment can be used to estimate the composition of microorganisms inhabiting that environment and to determine the types of biological activities taking place.

(i)   *Denaturation Gel Gradient Electrophoresis (DGGE)*

In DGGE, the phylogenetic marker sequence is amplified by PCR from the environmental metagenomic DNA using GC-clamped (chemiclamp) primers and the amplicons electrophoresed in a denaturing acrylamide gel (Schäffer and Muyzer, 2001). The concentrations of denaturants within the gel increase horizontally, parallel

to the electric field. As the DNA products move towards the cathode they encounter denaturants, which cause the DNA to melt and separate (Figure 1.15). Theoretically, each bacterial species will be represented by a unique band on the gel, the intensity of which is proportional to the abundance of that particular species in the whole community. Therefore data obtained can be used to make inferences about the abundance and diversity of the microbial population (Zoetendal *et al.*, 2001). For example, DGGE community profiles of endophytic bacterial communities in rice seeds showed the effect of vertical transmission and soil conditions on the composition of bacterial endophytes found in rice plants (Hardoim *et al.*, 2012). DGGE was also recently used to identify sorghum-associated endophytic bacteria in three geographical locations in South Africa (Ramond *et al.*, 2013).
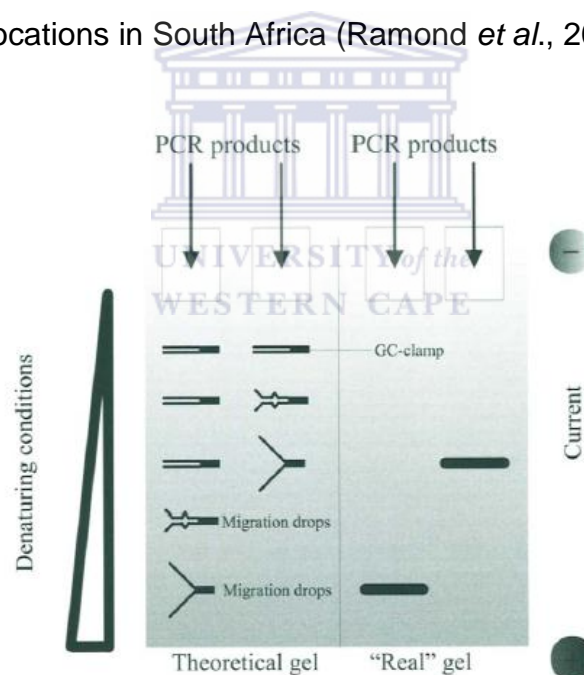
Figure 1.15. Schematic representation of migration of PCR amplicons through a DGGE gel (Zoetendal *et al.*, 2001).

A limitation of DGGE is that certain DNA fragments from different species could have the same melting patterns, thus resulting in comigration patterns within the gel. Inversely, some bacterial species have a heterologous mix of genes coding for the 16S rRNA and when they are all included in the PCR product mix, they could separate in the DGGE gel, appearing as more than one sample. This technique also has low resolution as only predominant bacterial groups are accessible (Zoetendal *et al.,* 2001).

### *(ii)	Terminal Restriction Fragment Length Polymorphism (t-RFLP)*

This fingerprinting technique was first developed by Avaniss-Aghajani *et al.* (1996) for the identification of mycobacteria in medical microbiology laboratories. Fluorescently-labelled primers are used in the amplification of a phylogenetic marker, which could be the bacterial 16S rRNA gene (Ding *et al.*, 2013) or a functional gene such as *nifH* (Sessitsch *et al.*, 2012), from metagenomic DNA. Purified amplicons are then digested using restriction enzymes and the labelled terminal restriction fragments (t-RFs) of different sizes are separated by capillary electrophoresis. Each t-RF represents an operational taxonomic unit (OTU).

T-RFLP has extensively been used extensively to study of endophytic bacterial communities. It is an effective technique in elucidating the diversity composition of bacterial communities in environmental samples. Ding *et al.* (2013) used t-RFLP to show (i) temporal shifts of endophytic bacterial community found in the leaves of *Asclepias viridis*, (ii) structural differences between endophytic communities found in

different plant species and (iii) effects of sampling site on the composition of endophytic communities (Figure 1.16).
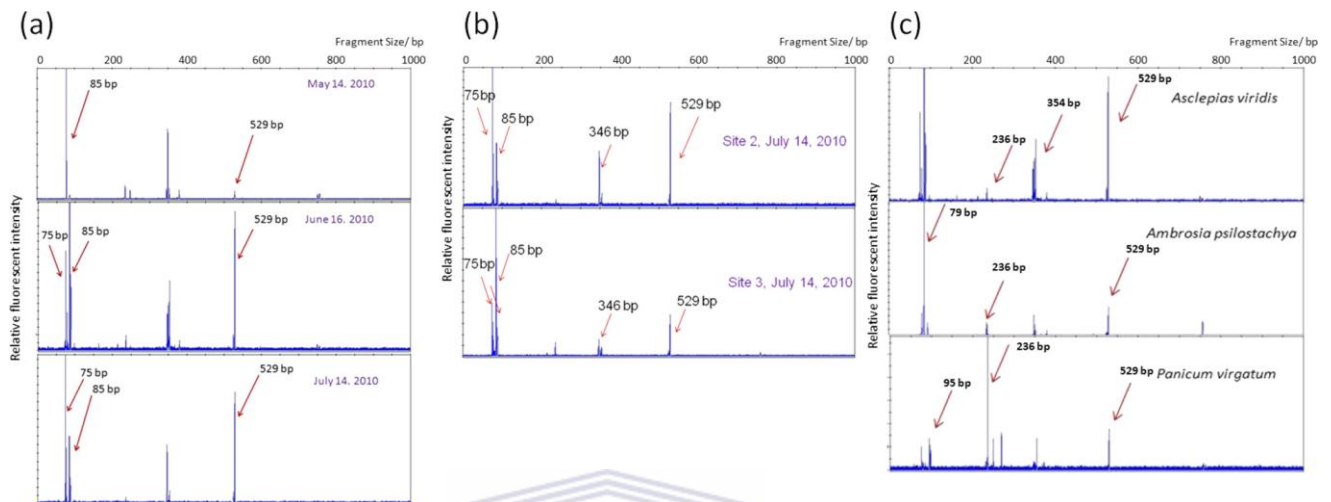
Figure 1.16. Comparisons of T-RFLP profiles of endophytic bacterial communities. Relative fluorescence intensity (normalized to the most intense peak) is plotted against length of the T-RF. T-RFLP profiles represented the bacterial species compositions, indicating the influences from multiple factors: (a) T-RFLP profiles from one tagged *Asclepias viridis* individual, samples of which were collected respectively on May 14th, June 16th and July 14th, 2010. (b) T-RFLP profiles from two *A. viridis* individuals respectively from Site 2 and Site 3, both collected on July 14th, 2010. (c) Selected T-RFLP profiles from 3 individuals respectively from *A. viridis*, *Ambrosia psilostachya* and *Panicum virgatum.*(Adapted from Ding *et al*. 2013).

Community profiling techniques are important in determining the structure of bacterial communities in complex environments. However, they are unable to resolve

the composition of these communities at the species level and rare constituents are also often overlooked in these analyses (Marschner *et al.*, 2005). Development of metasequencing tools that allow for parallel sequencing of phylogenetic markers from environmental samples has revolutionized microbial ecology studies, including studies of endophytic bacterial communities. The high resolution of these tools enables researchers to conduct in-depth structural and functional characterization of microbial communities (Marschner *et al.*, 2005).

c) Pyrosequencing

Pyrosequencing is a high-throughput sequence-by-synthesis technology where nucleotides are identified by the release of pyrophosphates when added to the template strand during DNA synthesis.
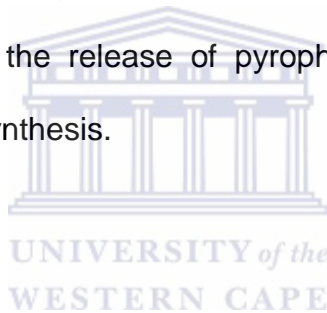
(i) Principles underlying pyrosequencing technology

During the sequencing process, addition of a nucleotide to an immobilised single-stranded DNA template (this could be a phylogenetic marker amplicon) triggers a series of enzymatic reactions that culminate in emission of light. The following steps/reactions outline the general principle of pyrosequencing (Ronalghi, 2001):

1. The (exo-) Klenow DNA polymerase, once attached to the primer, inserts a nucleotide (dNTP) to extend the complimentary strand. In this first reaction results in the release of PPi as follows.

$$(DNA)_n + dNTP \xrightarrow{\text{polymerase}} (DNA)_{n+1} + PPi$$

2. The PPi reacts with adenosine phosphosulphate (APS) during synthesis of ATP. This reaction is facilitated by ATP sulfurylase.

$$PPi \quad + \quad APS \xrightarrow{\text{ATP sulfurase}} \quad ATP \quad + \quad SO_4^{2-}$$

3. ATP reacts with luciferin in a luciferase mediated reaction to form a luciferin-luciferase-AMP complex, which in the presence of oxygen breaks up to produce oxyluciferin. Light is produced during this reaction, and it is captured by a CCD camera. This reaction confirms the insertion of a specific nucleotide, and this provides information about the sequence of the template strand.

$$\text{D-luciferin} + ATP \xrightarrow{\text{Luciferase}} \text{Luciferase -luciferin-AMP} \quad + \quad PPi$$

$$\text{Luciferase-luciferin-AMP} + O_2 \longrightarrow \text{Luciferase} + \text{oxyluceferin} + AMP + CO_2 + \text{light}$$
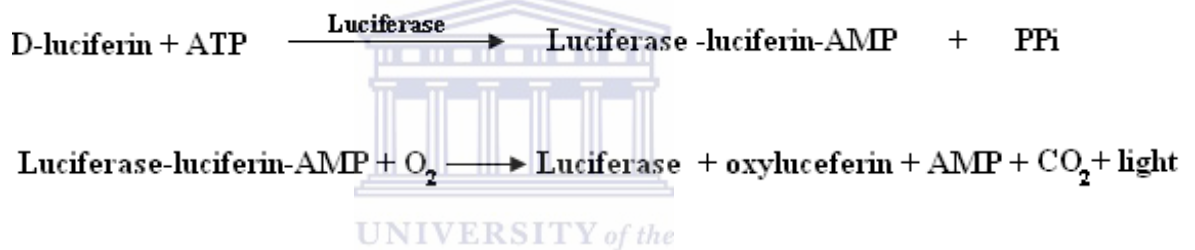
4. In the next reactions, the enzyme apyrase degrades the remaining nucleotides and ATP in the reaction medium in order to prevent interference of the next round of reactions.

$$ATP \xrightarrow{\text{Apyrase}} AMP + 2Pi$$

$$dNTP \xrightarrow{\text{Apyrase}} dNMP + 2Pi$$
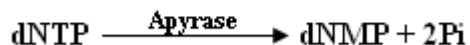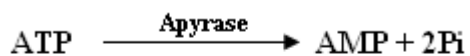
The massively-parallel nature of pyrosequencing allows for generation of great amounts of data in a relatively short period of time. In a single 24-hour run, the GS-FLX Titanium Series can generate up to 10^6 sequence reads of 400bp (www.454.roche.com). Pyrosequencing platforms continue to improve towards generation of longer sequence reads and increasing their throughput rates. Table 1.3

compares previous technologies (up to 2005) to the Roche FLX Titanium XL+ system that is currently used. This system can achieve read lengths of up 1000bp, compared to its predecessor, the GS FLX Titanium XLR70 which only managed up to 600bp (Roche http://454.com).

Table 1.3. Comparison of the current Genome Sequencer FLX+ Series to previously used pyrosequencing platforms (reviewed by Chan, 2005; www.454.roche.com)

Sequencing techology	Sample preparation	Read length (bp)	Throughput
Genome Sequencer FLX+ (current)	No Library	Up to 1000bp	700Mb/day per machine
454 Corp. pyrosequencing	No library	<100	2 Mbp/day per machine
Quake cycle extension	Library required	<50	10 bp/min per molecule
Polonies	No library	<50	10 bp/min per polony
Solexa cycle extension	Library required	<50	10 bp/min per molecule
Genovoxx cycle extension	Library required	<50	10 bp/min per molecule

(ii) Application of pyrosequencing technology in the study of endophytic bacterial communities

Pyrosequencing technology provides greater resolution than the previously discussed fingerprinting techniques in that individual microbial species can be accurately identified up to the species level (Charles, 2010).

Assemblages of endophytic bacteria associated with important plants and crops such as poplar trees (*Populus deltoides*) (Gottel *et al.*, 2011), saltbush species (*Atriplex canescens* and *Atriplex torreyi*) (Lucero *et al.*, 2011) and potato (İnceoğlu *et al.*, 2011) were elucidated through pyrosequencing of the bacterial 16S rRNA gene.

47

In whole-genome shotgun sequencing, the environmental DNA – in this case, plant metagenomic DNA – can be shredded and all fragments sequenced to provide data on an array of gene sequences available in that environment (Petrosino *et al*., 2005). Possible use of pyrosequencing in the study of plant diseases was shown in a deep sequencing metatranscriptomic study that revealed the presence of viruses and pathogenic bacteria and fungi in the tissues of leguminous soybean plants (Molina *et al*., 2012). This technology also enabled the inclusion of previously uncultured bacterial species in ecological studies and discovery of natural products from plant tissues (Berlec, 2012).

Pyrosequencing technology can also be applied in whole genome sequencing studies of important endophytic bacteria. For example, the genome of *Variovorax paradoxus*, an endophytic bacterium capable of degrading various soil pollutants including chemical fertilizers and pesticides, was recently decoded using pyrosequencing (Han *et al*., 2011). In this study, 6279 proteins were predicted indicating the great metabolic potential of this microorganism that is yet to be explored. This is also interesting as a sequenced bacterial genome constitutes a reliable reference for other studies pertaining to that particular organism, and can also serve as model organisms in the study of other species (Krause *et al*., 2007; Bertalan *et al*., 2009; Han *et al*., 2011).

1.3. Motivation of study

This study aims to investigate the diversity of endophytic bacterial communities associated with two important cereal crops farmed in South Africa; namely sorghum (*Sorghum bicolor* L. Moench) and pearl millet (*Pennisetum glaucum* L.), with an emphasis on plant growth promoting endophytes with biofertilizer and/or biocontrol potential. The effect of DNA extraction protocols on the metagenomic DNA quality and its representation of native endophytic communities is also compared using two plants: sorghum, a monocotyledonous plant and groundnut (*Arachis hypogeae*), a dicotyledonous legume.

1.3.1. Exploitation of plant-microbe interactions for production of healthy crop

The development of alternative agricultural strategies for production of high-yield, healthy agricultural crops is influenced by the challenges originating from food shortage problems in the African continent. Over 50% of the African population live in poverty or poor nutrition, and the widespread poverty and malnutrition in most African countries has been closely linked to their low productivity levels, vulnerability to shocks such the HIV/AIDS pandemics and high illiteracy levels (Stige *et al*., 2006). These consequences further affect the crippling economies, and it becomes a vicious cycle. African initiatives such as the Comprehensive Africa Agriculture Development Programme by the New Partnership for African Development (NEPAD) and Millenium Development Goals (MDGs) have placed the development of

agricultural land and produce at the core of their strategies in order to address these challenges (Skoet *et al.*, 2004).

This study falls within the scope of agricultural and food management as it is aimed to study and understand the diversity of endophytic microorganisms associated with food crops, with focus on potential development of environmentally-friendly novel biotechnological approaches.

1.3.1.1. Sorghum and pearl millet

Sorghum (*Sorghum bicolor* L.) and pearl millet (*Pennisetum glaucum* L.) (Figure 1.17) are drought-tolerant monocotyledonous plants belonging to the same family, Poaceae, therefore they are well-adapted to the semi-arid and sub-tropical regions of Africa, a continent that actually contributes to approximately 50% of the global production of these commercially important crops (Belton and Taylor 2004).

These grains are rich in carbohydrates and protein (Table 1.4) and constitute staple foods in most African countries where they are consumed as cereal or as an ingredient in traditional dishes such as porridge, bread and cakes. Sorghums and millets are also used as substrates in various fermentation processes including home- and commercial beer brewing (Agu *et al.*, 1998). Sorghum grains and molasses are also used in bioethanol production (Sheorain *et al.*, 2000; Ai *et al.*, 2011).
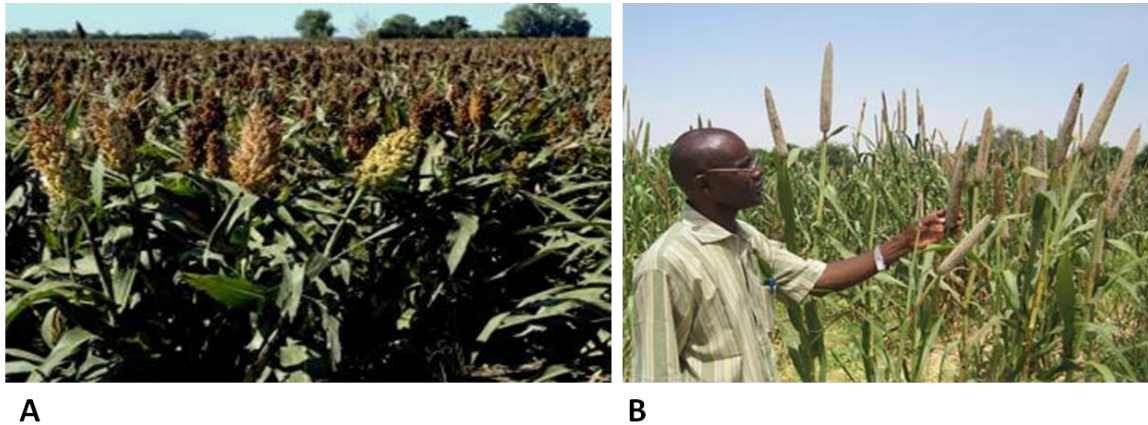
Figure 1.17. Mature sorghum (A) and pearl millet (B) crop (Courtesy of Independent Online and ICRISAT, 2013).

Table 1.4. Nutritional information of sorghum, millet and peanut (dietary proximate per 100g of seed/grain). (USDA, National Nutrient Database for Standard Reference)

Food Type	Water (g)	Energy (kcal)	Protein (g)	Lipids (g)	Carbohydrates (g)	Fibre (g)
Sorghum	9.2	339	11.3	3.3	74.63	6.3
Millet (Raw)	8.67	378	11.02	4.22	72.85	8.5
Peanut	6.5	567	25.8	49.24	16.13	8.5

1.3.1.2. Groundnut

Groundnut (*Arachis hypogeae*, L.) (Figure 1.18), also known as peanut, is a legume plant that belongs to the Fabaceae family. Peanuts are commercially produced in over 90 countries, providing a global annual production of 36.45 million tons and yield of 1524 kg/ha in 2009 (FAOSTATS, 2011). Asia is the leading producer peanuts, and accounts for 64% of global production. African countries such as Nigeria, Sudan, Senegal, Chad, Ghana, Congo and Niger are also important exporters of peanuts. South Africa only accounts for less than 2% of the continental

groundnut production, and is a major importer of this food crop. Groundnut farming is mostly done by smallholder farmers, and it provides employment opportunities for nearby communities (I-Life, 2005).

Figure 1.18. A harvested groundnut crop showing edible seeds growing from the roots (Adapted from ICRISAT, 2013).
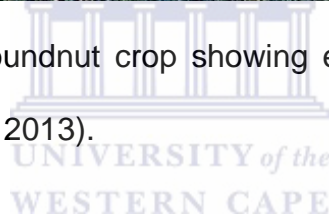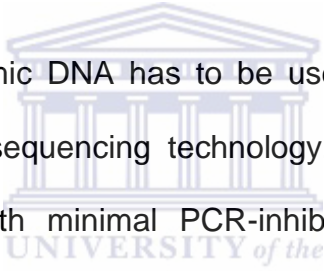
Groundnut kernel is an important oilseed that is also eaten as a snack or as an ingredient in various foods. Peanut kernels are a popular snack amongst kids and adults, and are edible when roasted or unroasted. These seeds are highly rich in proteins, carbohydrates, lipids, micronutrients and vitamins, and because of their affordability, they were nicknamed "the poor man's snack" (Table 1.4) (Settaluri *et al.*, 2012).

1.3.2. Aims and objectives

The aim of this study is to investigate the diversity of endophytic communities associated with *Sorghum bicolor* (L.) Moench (sorghum) and *Pennisetum glaucum* (pearl millet) using Next Generation Sequencing, and three objectives have been identified to achieve it:

1. <u>To determine the most efficient protocol for extraction of good-quality metagenomic DNA from plant tissues to study endophytic bacterial communities.</u>

    High-quality metagenomic DNA has to be used in this study of endophytic bacteria because pyrosequencing technology, as a PCR-based technique, requires pure DNA with minimal PCR-inhibiting contaminants. Also, the protocol used has to be able to access the genomic material greater majority of the endophytic bacterial community in order to minimise potential underestimation of the community. In this study seven DNA extraction protocols (5 commercial kits and 2 classical protocols) will be used to extract metagenomic DNA from root and stem tissues of sorghum and groundnut plants. The efficiency of the different protocols will be compared on the basis of yield, purity and quality for DNA in use in PCR reactions. Diversity of endophytic bacterial communities retrieved by each protocol will be determined and compared with t-RFLP.

2. <u>To determine the endophytic bacterial community structures and assemblages associated with the root and stem tissues of two staple food crops</u>

<u>farmed in South Africa (sorghum and pearl millet) using 454 pyrosequencing technology.</u>

Previous t-TFLP-based studies at IMBM have indicated that there could be a core sorghum-associated rhizospheric bacterial community that is independent of environmental or edaphic factors (Ramond *et al.*, 2013). However, findings in this study were inconclusive regarding the effects of environmental factors on the endophytic communities, which were also shown to have low diversity. The current study is a continuation of this endeavour.

High-throughput pyrosequencing technology is employed to improve the sensitivity and resolution in the detection and identification endophytic bacteria from plant tissues. Sorghum is also compared to a closely related plant, pearl millet, to determine shared phylotypes between the two plants, and potential for recovering phylotypes exclusive to either species. These communities are compared on the basis of their diversity, biofertilizer/biocontrol potential, and potential for other industrial applications.

CHAPTER 2

Materials and Methods

CONTENTS

2.1. Chemicals, reagents, media and kits

Table 2.1 outlines an exhaustive list of chemicals, reagents, kits and enzymes used in this study. Buffer compositions are shown in Table 2.2.

Table 2.1. Chemicals and reagents.

Reagent/Kit	Supplier
6X DNA loading dye	Fermentas
24:1 chloroform isoamyl alcohol	Sigma
Agarose	Lonza
β-mercaptoethanol	Merck
Hexadecyltrimethylammonium bromide (CTAB)	Sigma
Ethylene diamine tetra acetic acid (EDTA)	Merck
Ethanol	Merck
Ethidium bromide	Sigma
GeneJET Plant Genomic DNA Purification Kit	Fermentas
Glacial acetic acid	Merck
Glucose	BDH
Hydrochloric Acid (HCl)	Merck
Illustra GFX™ PCR DNA and Gel Purification Kit	GE Healthcare
Isopropanol	Merck
Lysozyme	Sigma
MoBio PowerPlant Pro® DNA Isolation Kit	Optima Scientific cc
MoBio PowerSoil™ DNA Purification Kit	Optima Scientific cc
MoBio UltraClean® Soil DNA Isolation Kit	Optima Scientific cc
Nutrient Agar	Merck
Phenol	Sigma
Polyvinylpyrrolidone (PVPP)	Sigma
Proteinase K	Fermentas
R2A Agar	Merck
Restriction Enzymes (*Hae*III, *Hind*III, *Pst*I)	Fermentas
RNase A	Fermentas
Sodium chloride (NaCl)	Merck
Sodium dodecyl sulphate (SDS)	Promega
Sodium hypochlorite	Kimix
Tris (Tris[hydroxymethyl] aminoethane)	BDH
Qiagen MinElute® PCR Purification Kit	Kapa Biosystems
Qiagen DNeasy[R] Plant Mini Kit and Fermentas	Whitehead Scientific

Table 2.2. Buffers used in this study.

Buffer	Components
Lysozyme buffer	25mM TRIS-HCl
	50mM glucose
	10mM EDTA
	25mg.mL^{-1} lysozyme
TE buffer (pH 8)	10mM Tris-HCl
	1mM EDTA
Double strength CTAB buffer	100mM Tris-HCl
	1.2M NaCl
	20mM EDTA
	2% CTAB
	0.2% β-mercaptoethanol
1X TAE buffer (pH8)	40 mM Tris acetate
	1 mM EDTA
	0.2mM glacial acetic acid
Phosphate Buffer solution (PBS)	140mM NaCl
	2.5mM KCl
	10mM Na$_2$HPO4.2H$_2$O
	1.5mM KH$_2$PO$_4$

2.2. Plant collection

The studied plants, i.e. sorghum, pearl millet and groundnut were obtained from experimental plots of the Agricultural Research Council (ARC) farm situated in Potchefstroom (North West Province, South Africa) (Figure 2.1). Five mature and healthy plants were collected for each plant type, following a random sampling technique and using sterilized gardening forks and shovels. Leaves were removed with ethanol-sterilised side-cutters. Stems were cut into approximately 15cm long pieces and placed in sterile bags. The roots were shaken in sterile collection bags to collect the rhizospheric soil prior to storage in separate sterile bags. All samples were immediately placed on ice and transported to the Institute for Microbial Biotechnology and Metagenomics (IMBM), where they were stored at -80°C prior to processing.

Figure 2.1. Location of Potchefstroom on the South African map.

The growing conditions and state of the different plants at the time of collection (as provided by the ARC) are hereby detailed, and locations of individual samples are shown in Table 2.3.
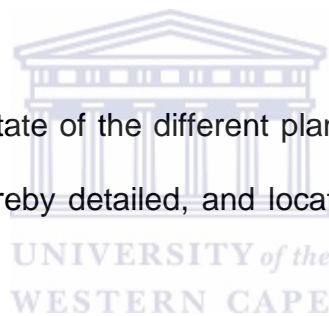
a. Sorghum (*Sorghum bicolour* (L.) Moench)

At the time of sampling, sorghum cultivars were 16 weeks old and therefore considered mature. This crop has been planted on the sampled field annually for a period of four years. Water was primarily supplied via rainfall events. In the case of low rainfall levels (i.e. below the expected annual average [~300-320mm]) (South African Weather Services, 2012), irrigation was supplied. The soil was fertilized with "3:2:1 (32) + ZN" fertiliser (a slow-release fertiliser that contains zinc and a 32% mixture of nitrogen (N), phosphorus (P) and potassium (K) in the ratio 3:2:1 at a rate of 150kg N/ha.

LAN 28, a fertiliser containing 28% total nitrogen with 50% ammonium and 50% nitrate, was applied at a rate of 100kg/ha when plants were at knee length. The soil was also treated with the insecticide, Kombat® (Kombat, South Africa, www.kombat.co.za), for control of stem-borers.

b. Pearl millet (*Pennisetum glaucum*)

Pearl millet plants were 14 weeks old at the time of sampling. They were obtained from a field where rotational farming is practised, alternating between pearl millet and sunflower. Exact period for this practice was not confirmed, but it is believed to exceed 5 years (personal conversation with Dr. Nemera Shargie, ARC). The current cultivar, Common, was supplied to the ARC by a seed biotechnology company, Agricol (South Africa; www.agricol.co.za). The field was ploughed before planting and the crop was grown tillage-free without supplementary irrigation despite the below average rainfall. Fertiliser 3:2:1 (25), which contains 25% mixture of N, P and K was applied at 105kg/ha at planting. At the time of sampling, plant growth appeared to be underdeveloped, which could be attributed to drought conditions and growth of weeds on the plot. In some sections of the plot an aphid infestation was observed and therefore no plants from these areas were sampled.

c. Groundnut (*Arachis villosulicarpa*)

Five months old groundnut plants were obtained from a rotational field where maize and groundnut planting alternated annually. Prior to planting, the field was treated with a series of deep plough, disc and corn skiller, and the herbicide, Roundup ® (Monsanto SA, Bryanston, South Africa; www.monsanto.co.za) was applied. At

flowering stage, the soil was supplemented with calcium fertiliser (gypsum) at a rate of 1ton/ha for improved pod development. A fungicide, Punch® (Du Pont, South Africa; www.dupont.com) was also sprayed twice to control foliar disease.

Table 2.3. Original locations of plants used in this study.

Crop	Sample	GPS Coordinates
Sorghum	S1	S26^{O}44302' E027^{O}05584'
	S2	S26^{O}44312' E027^{O}05588'
	S3	S26^{O}44321' E027^{O}05581'
	S4	S26^{O}44315' E027^{O}05584'
	S5	S26^{O}44302' E027^{O}05572'
Pearl millet	M1	S26^{O}44144' E027^{O}04649'
	M2	S26^{O}44156' E027^{O}04651'
	M3	S26^{O}44159' E027^{O}04642'
	M4	S26^{O}44151' E027^{O}04642'
	M5	S26^{O}44167' E027^{O}04642'
Groundnut	G1	S26^{O}44276' E027^{O}03606'
	G2	S26^{O}44274' E027^{O}03600'
	G3	S26^{O}44274' E027^{O}03601'
	G4	S26^{O}44267' E027^{O}03600'
	G5	S26^{O}44268' E027^{O}04605'

2.3. Plant tissue preparation and sterilisation

The plant organs were sterilised using a modified protocol designed by Mendes *et al.*, (2007). The roots and stems of each plant species (sorghum, pearl millet and groundnut) were separately washed in autoclaved double-distilled water until all residual soil was removed from their surfaces. Plant organs were immerged in 500mL 1X phosphate buffer solution (PBS) for 1.5 hours, shaking at 100rpm. The samples were then sequentially washed by shaking in (i) 70% ethanol for 10 minutes, (ii) 2.5% sodium hypochlorite solution for 20 minutes and (iii) rinsed five

times in autoclaved double-distilled water for 2 minutes. To confirm sterility, 100μL of the last wash water was inoculated on nutrient agar (NA) and R2A agar plates (in triplicate) and incubated at room temperature for three days. The plants were stored in the last wash water at 4°C during these 3 days. Where no colony growth was observed, the sterilisation process was considered successful. Where colonies were observed, the sterilisation process was repeated. A repeat sterilisation was conducted at least once per sample in this study.

Sterilised plant tissue was aseptically ground to a fine powder in liquid nitrogen using autoclaved pestle and mortar. Ground tissue powder aliquots (100mg) were then stored at -80°C.

2.4. Molecular Biology

2.4.1. Metagenomic DNA extractions

Seven different protocols were used to extract metagenomic DNA from 0.1g or 0.3g root or stem tissues of sorghum, pearl millet and groundnut and compared.

Two of these were classical DNA extraction protocols, being either SDS-based (Zhou *et al.*, 1996) or CTAB-based (Murray and Thompson, 1980). The five remaining were commercial kits previously used to study endophytic communities (Green *et. al.*, 1999; Drabkova *et. al.*, 2002; Krechel *et al.*, 2002; West *et al.*, 2010): MoBio PowerPlant Pro® DNA Isolation Kit, Qiagen DNeasy^{R} Plant Mini Kit and

Fermentas GeneJET Plant Genomic DNA Purification Kit, MoBio PowerSoil™ DNA Purification Kit and MoBio UltraClean® Soil DNA Isolation Kit. All kit-based DNA extraction protocols were performed according to the manufacturer's instructions, with the exception that starting plant material quantity was always 0.1g or 0.3g and a 50µL final elution was performed. All extractions were carried out in triplicate. Classical protocols are described in detail below:

SDS-based protocol (Modified protocol: Zhou *et al*, 1996):

Five hundred microlitres of lysozyme buffer and RNase A (final concentration 50µg.mL^{-1}) were added to 0.1g or 0.3g ground plant tissue and mixed. The mixtures were incubated at 37°C for 1 hour, and then treated with Proteinase K (final concentration 1mg.mL^{-1}) at 37°C for 1 hour. SDS (1% final concentration) was added and mixed by flicking and inverting the tubes ten times, and mixtures were incubated at 65°C for 30 minutes. Tubes were centrifuged (14000rcf, 2 minutes) and the supernatants collected into new tubes. Equal volume phenol was added to each tube and mixed by inversion. Top aqueous phase containing DNA was collected after centrifugation (10000rcf, 1 minute) and the bottom layer with organic phenol was discarded. The phenol extraction was repeated once. Equal volume 24:1 (v/v) chloroform/isoamyl alcohol solution was added to each tube and mixed by inversion. Top aqueous layer was collected and transferred to a new tube after centrifugation (10000rcf, 10 minutes). The tubes were placed on ice and equal volume ice-cold isopropanol was added, followed by incubation at 4°C for 20 minutes. The tubes were centrifuged (10000rcf, 5 minutes) to recover metagenomic DNA and the isopropanol was discarded. DNA pellets were air dried under the laminar flow

cabinet and then washed twice with 250μL 70% ethanol, which was eluted after centrifugation (10000rcf, 5 minutes). The DNA pellets were allowed to dry and were then resuspended in 50μL autoclaved TE buffer and stored at -20°C.

CTAB-based protocol (Modified protocol: Murray and Thompson, 1980):

Double strength CTAB buffer (700μL) was added to ground plant tissue. The mixture was vortexed for 20 seconds (maximum speed), followed by incubation at 65°C for 1 hour and addition of 600μL 24:1 (v/v) chloroform/isoamyl alcohol solution. The tubes were mixed by inversion and centrifuged (12000rpm, 5 minutes). Equal volume of ice-cold isopropanol and RNase A (final concentration 50μg.mL^{-1}) were added to the supernatant in a clean tube and mixed by inversion. The tubes were incubated at room temperature for 20 minutes and then centrifuged (12000rpm, 5 minutes). The supernatant was discarded, and the pellets were allowed to air dry in a laminar flow hood cabinet. DNA pellets were washed twice with 250μL 70% ethanol, which was eluted following centrifugation (12000rpm, 5 minutes). The DNA pellets were allowed to air dry in a laminar flow cabinet and then resuspended in 50μL TE buffer before storage at -20°C.

2.4.2. DNA quantification and purity

Metagenomic DNA and PCR product concentrations were estimated using the Nanodrop ND-1000 Spectrophotometer (NanoDrop Technologies, USA). One A_{260} OD unit is equivalent to 50ng.mL^{-1} double stranded DNA. DNA was considered pure

where $1.7 \leq A_{260}/A_{280} \leq 1.9$. Final quantification of pyrosequencing amplicons was performed using the Qubit® Flourometer (Invitrogen).

2.4.3. Gel electrophoresis

Metagenomic DNA and PCR products were separated by electrophoresis on 0.8% and 1.5% agarose gel, respectively. Five microlitres of DNA sample was mixed with $1\mu L$ 6X loading dye (DNA tracking dye) and loaded on the agarose gel containing $50\mu g.mL^{-1}$ ethidium bromide for staining and visualisation. Electrophoresis were performed in 1X TAE buffer at 80V, 2.5 hours for total extracted DNA and 30 minutes for PCR products. DNA sizes were determined by comparing band migration to migration of DNA molecular weight marker bands (i.e. lambda DNA cut with *Hind*III or *Pst*I restriction enzyme). Gels were visualised under ultraviolet illumination and photographed with a digital imaging system (AlphaImager 2000, Alpha Inotech, San Leandro, CA).

2.4.4. Polymerase chain reaction (PCR)

The Labnet MultiGene™ Gradient PCR Thermal Cycler (Labnet International, Inc.) was used for all PCR reactions. PCR conditions are outlined in Table 2.4. The E9F-U1510R primer set was used to amplify the bacterial 16S rRNA gene from plant metagenomic DNA. Initial tests were to determine the optimal template DNA concentration for PCR amplification of the bacterial 16S rRNA gene. These reactions

were carried out in 50µL volumes containing 1X DreamTaq Buffer, 0.2mM each dNTP, 0.5M each primer, template DNA (1ng, 5ng or 10ng), 0.3µL DreamTaq DNA polymerase (Fermentas, Lithuania) and deionised nuclease-free water. The same primer set was used during amplification of the 16S rRNA gene for terminal restriction fragment polymorphism analysis, only here, the forward primer (E9F), was labelled at the 5' end with fluorescent dye fluorescein amidite (FAM).

Table 2.4. Primer combination and respective cycling conditions used in this study.

Primer	Sequence (5' to 3')	Amplification cycle	References
E9F	GAGTTTGATCCTGGCTCAG	$95^{o}C$, 4min; 30 cycles of $95^{o}C$ for 30s, $52^{o}C$ for 30s and $72^{o}C$ for 105s; $72^{o}C$, 10min.	Turner et al., 1999
U1510R	GGTTACCTTGTTACGACTT		
8F	CCATCTCATCCCTGCGTGTCTCCGAC	$95^{o}C$, 2min; 25 cycles of $98^{o}C$ for 20s, $75^{o}C$ for 15s and $72^{o}C$ for 90s; $72^{o}C$, 10min.	(Turner *et al.*, 1999; Muyzer *et al.*, 1993)
518R	ATTACCGCGGCTGCTGG		

Amplification of the bacterial 16S rRNA gene for pyrosequencing analysis was carried out using the 8F-518R primer set in 25µL volumes containing 1X Phusion HF Buffer, 200µM each dNTP, 0.5µM each primer, 50ng template DNA, 0.02U/µL Phusion High Fidelity DNA Polymerase (Fermentas, Lithuania) and deionised nuclease-free water. These primers were modified by annealing them to sample specific extended multiplex identifier (MID) adaptor sequences for the Genome Sequencer (GS) Titanium FLX Chemistry (Roche) as shown in Table 2.5.

Table 2.5. Identification of bacterial 16S rRNA sequence source using unique MID sequences.

Plant Tissue	MID sequence (5'-3')	Adaptor sequence (5'-3')	Tag Sequence (5' -3')
Sorghum stem	ACGCTCGACA	CCATCTCATCCCTGCGTGTCTCCGAC	TCAG
Sorghum root	CGTGTCTCTA	CCATCTCATCCCTGCGTGTCTCCGAC	TCAG
Pearl millet stem	TAGTATCAGC	CCATCTCATCCCTGCGTGTCTCCGAC	TCAG
Pearl millet root	TCTCTATGCG	CCATCTCATCCCTGCGTGTCTCCGAC	TCAG

2.4.5. DNA purification

(a)   Metagenomic DNA purification using polyvinylpyrrolidone (PVPP)

Metagenomic DNA samples were cleaned using self-made PVPP spin columns based on a procedure described by Berthelet *et al*. (1996). A spin column was constructed by placing a 20μL filter tip (cut 2mm underneath the filter) inside a 0.6mL tube with the bottom and the cap cut off. The column was placed inside a 1.5mL eppendorf tube. PVPP suspensions in TE buffer (150μL, 100g.L^{-1}) were loaded twice into spin columns and centrifuged (1500rpm, 2 minutes). The PVPP residues were washed twice with 150μL TE buffer and centrifuged (1500rpm, 2 minutes). The supernatant was discarded and the residues were dried by centrifugation (3000rpm, 10 minutes). The spin column was put into a clean collection tube. Metagenomic DNA samples were aliquoted directly onto the PVPP in the spin column and incubated at room temperature for 1 minute. The DNA was then eluted by centrifugation (3000rpm, 5 minutes; 5000rpm, 10 minutes).

(b)   PCR amplicon purification using commercial kits

The Illustra GFX™ PCR DNA and Gel Purification Kit (GE Healthcare, UK) was used to purify PCR amplicons prior to t-RFLP analysis. Amplicons prepared for pyrosequencing were purified using the Qiagen MinElute® PCR Purification Kit. Both kits were used according to manufacturer's specifications.

2.4.6. Terminal Restriction Fragment Polymorphism (t-RFLP)

The bacterial 16S rRNA gene was amplified from sorghum and groundnut metagenomic DNA as described (Section 2.4.4). Purified PCR amplicons (200ng per reaction) were digested overnight at 37°C using the restriction enzyme *Hae*III restriction enzyme. The lengths of fluorescently labelled terminally restricted fragments (t-RFs) were determined on an Applied Biosystems Genetic Analyzer sequencer at the Central Analytical Facility of the University of Stellenbosch, and using the internal size standard marker, ROX1.1 (sizes in bp: 75, 100, 139, 150, 160, 200, 300, 340, 350, 400, 450, 490, 500, 583, 683, 782, 932, 991, 1121). T-RFLP patterns were analysed using the Peak Scanner™ Software Version 1.0 (Applied Biosystems). Valid peaks (between 35 and 1000) were identified and analysed using the online T-REX software (http://trex.biohpc.org/; Culman *et al*., 2009). T-RFs were characterised by peak height and aligned to create an operational taxonomic unit (OTU) data matrix. The term OTU refers to an individual t-RF. Theoretically, one t-RF comprises of one bacterial ribotype (Blackwood *et al*., 2007). However, in the case where the restriction enzyme recognition site lies at the same base position on

the 16S rRNA gene from different bacterial species, the generated t-RF would comprise of more than one bacterial ribotype (Nocker *et al.*, 2007).

The OTU matrix was analysed with the software Primer 6, Version 6.1.11. (Primer E, Plymouth, UK). Diversity indices were calculated for each sample using the diverse function in Primer 6. Standardised t-RFLP profiles were used to calculate the Bray-Curtis similarity coefficients (Bray and Curtis, 1957) between samples, which were then used to create similarity matrices of presence-absence transformed data. Similarity matrices were used to construct non-metric multidimensional (nMDS) plots, which are ordinations of sample communities based on their relative similarities, i.e. the distances between two points reflects the degree of similarity between microbial community profiles (Sherphard, 1962; Clarke & Warwick, 2001).
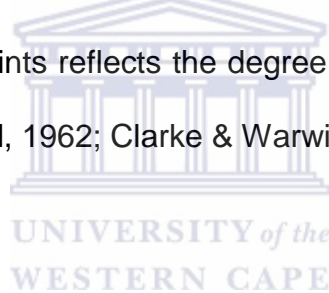
A selection of predominant OTUs was assigned to known bacterial taxa by *in silico* digestion of the bacterial 16S rRNA gene using the web-based tool MiCA (Microbial Community Analysis; Shyu *et al.*, 2007) with the "RDP (R10, U27) 700, 829 Good Quality (>1200) Bacterial" database. A ±3 bp size margin was permitted to account for potential differences between real and predicted T-RFs due to possible T-RF drifts.

2.4.7. Pyrosequencing of the bacterial 16S rRNA gene

For pyrosequencing, metagenomic DNA was extracted from sorghum and pearl millet root and stem tissues using the SDS-based classical protocol. Metagenomic DNA samples (four per plant tissue type) of good quality and possessing high endophytic bacterial diversity (determined by t-RFLP) were selected for this experiment. The V2-V3 region of the 16S rRNA gene was amplified from these samples using the 8F-518R primer set. Five PCR reactions were prepared per sample, and PCR products from each tissue type were pooled during the purification process and quantified. The amplicons were submitted to the Next Generation Sequencing Facility at the University of the Western Cape for pyrosequencing on the Roche 454 GS Junior System.

2.5. Data Analysis

Statistical tests for DNA yield and quality as well as diversity indices were performed using the software Sigma-Plot, Version 11.0. (Systat Software, Inc.). Two-way analysis of variance (ANOVA) was used to compare yield and purity of DNA extracted from six plant tissues (root and stem of sorghum, pearl millet and groundnut) using five DNA extraction protocols. Normality tests were performed on the data following the method of Kolmogorov-Smirnov, with Lillifor's correction (Justel *et al.*, 1997). Data that did not pass the normality test was compared using the Holm-Sidak test that ranks the ordinal numbers and compares the median of the samples (Holm, 1979). Paired T-tests were conducted to compare differences in

DNA yield at 0.1g and 0.3g starting plant material within individual tissues. Where normality criteria were not met for a paired T-test, the Wilcoxon Signed Rank test (Wilcoxon, 1945) was used.

Outputs from the 454 GS Junior software included a quality file (QUAL format) with information on the sequencing process and a metadata file (FASTA format) containing raw 16S rRNA sequences. The two files were processed using the CloVR-16S pipeline version 1.1., which comprises of a suite of phylogenetic tools (Figure 2.2; Angiuoli *et al.*, 2011). Preprocessing, processing and analysis of data were done using modules in QIIME (Quantitative Insights into Microbial Ecology http://qiime.org), R (http://www.R-project.org/) and MOTHUR (Schloss *et al.*, 2009).

Quality assessment and filtering of raw sequences was performed using Prinseq (Schmieder and Edwards 2011). Multiplexed reads were split and assigned to samples based on their MID sequences using a Python script. Sequences were then trimmed and filtered to include only good-quality sequences of 200-470bp. Sequences with high ambiguous base (N) occurrence and poly-A/T tails were removed. *De novo* chimera detection and OTU picking were performed with UCHIME and UCLUST, respectively (Edgar, 2010, Edgar *et al.*, 2011). In this study, an OTU is defined as a cluster of sequences (Floyd *et al.*, 2002), delimited at ≥97% sequence similarity. The representative sequences were aligned and used to create an OTU distance matrix with PyNAST (Caporaso *et al.*, 2010). The distance matrix was converted to a phylogenetic tree with FastTree (Price *et al.*, 2009). Taxonomy was assigned based on the Greengenes taxonomy and a Greengenes reference

database (version 12.10) (McDonald *et al.*, 2012), using the RDP Classifier (version 2.2) (Werner *et al.*, 2012).

The OTU table was used to calculate the alpha diversity (within sample diversity) and generate rarefaction plots with QIIME. To create a rarefaction curve, the whole community is subsampled repeatedly at fixed sample size increments and diversity metrics are calculated at each interval and plotted as a graph (Gotelli and Colwell, 2011). In this study, the Chao1 diversity index was calculated as a measure of true species diversity using the formula, $S_1 = S_{obs}(F_1^2/2F_2)$, where S_{obs} is the number of species observed, F_1 is the number of singletons (occur once) and F_2 is the number of doubletons (Gotelli and Colwell, 2011). Simpson index (1-λ) was calculated to measure community evenness.
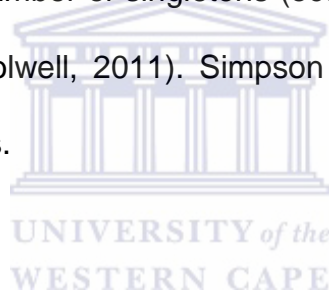
The taxonomic predictions for each OTU were used to create heatmaps and bar charts to reflect the distribution of phylotypes within each sample. UniFrac was used to determine beta-diversity, which is an assessment of differences between samples (Lozupone and Knight, 2005; Lozupone *et al.*, 2006; 2007). In this case, the phylogenetic tree is used to create a distance matrix where "distance" or dissimilarity between each pair of community samples is calculated, and this information is used to create a Principal Coordinate Analysis (PCoA) plot.

UniFrac was then used to determine whether there was a significant difference in the structure of endophytic bacterial communities retrieved from the four tissues. Briefly,

UniFrac measures the distance between each pair of environments as a fraction of the total branch length in a phylogenetic tree, leading to sequences of one environment (Lozupone and Knight, 2005). To compare pairs of environments, the UniFrac value is calculated for each pair and a distance matrix is created. The distance matrix was then geometrically converted into a Principal Coordinates Analysis (PCoA) plot, which is an ordination of bacterial communities in space according to their phylogenetic similarities. A PCoA plot consists of a series of orthogonal axes along which the amount of variation between environments is maximised. Pairs of environments were also subjected to a statistical UniFrac test (F-test) to test for significance of difference at 95% confidence level. All analyses were carried out using the unweighted UniFrac metric for qualitative comparisons of β-diversities not affected by individual sequence abundance.
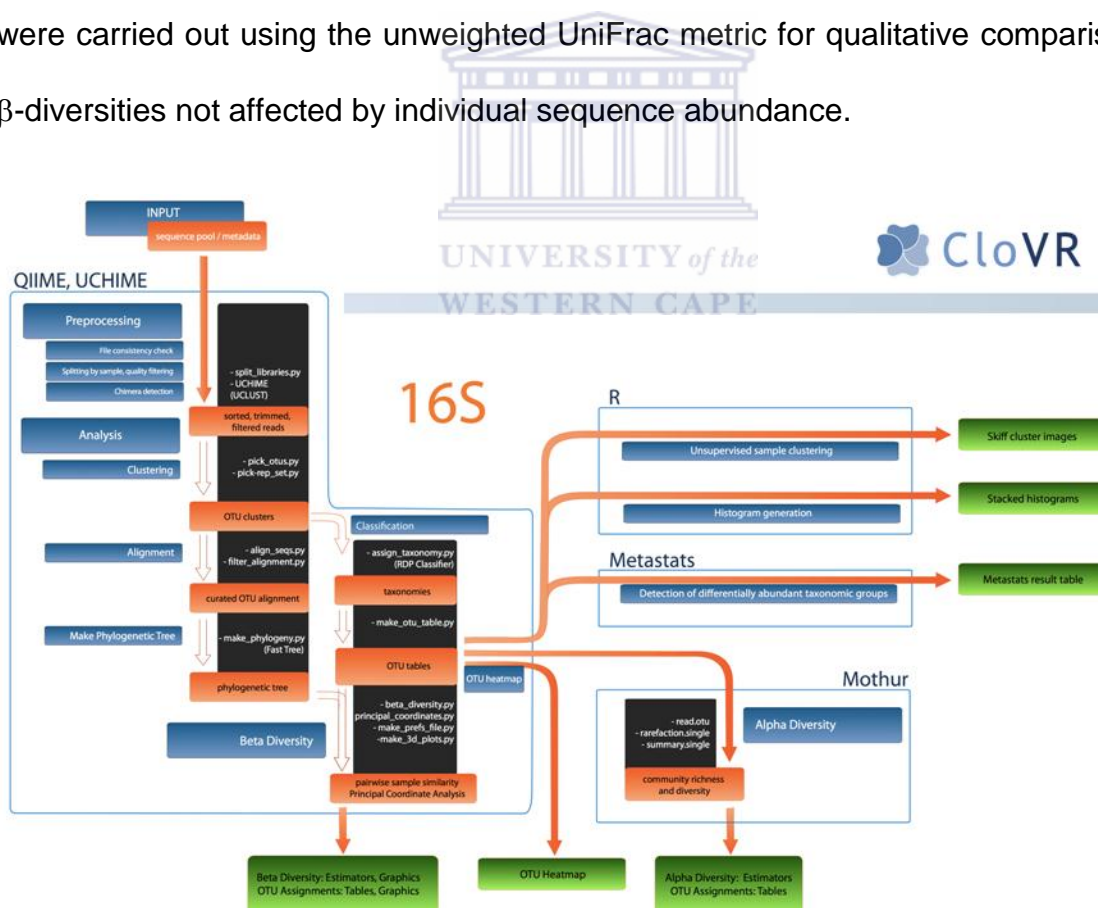
Figure 2.2. A workflow outline of the data analysis process followed when using the CloVR pipeline.

CHAPTER 3

Effects of DNA Extraction Protocols on

Endophytic Bacterial Community

Diversities

CONTENTS

3.1. Introduction

Metagenomic approaches are integral to ecological studies of microbial communities found in complex environments such as water (Gomez-Alvarez *et al.*, 2012), soils (Monteiro *et al.*, 2009; Ramond *et al.*, 2012), faeces (Salonen *et al.*, 2010), hot springs (Sharma *et al.*, 2007) as well as plant tissues (Gottel *et al.*, 2011). In these culture-independent studies, genetic material (DNA and/or RNA) is extracted directly from environmental habitats to retrieve information on their indigenous microbial community structure and its functional potentials (Marschner *et al.*, 2005).

One of the primary challenges faced in culture-independent studies of plant-associated microorganisms – as with other complex environments – is the retrieval of good-quality metagenomic DNA that can further be analysed, notably via PCR-based community profiling techniques (e.g. t-RFLP and DGGE) (Sessitsch *et al.*, 2002) and next generation sequencing (NGS) (Demeke and Jenkins, 2010). The DNA extraction process employed in a metagenomic study should therefore, (i) ensure lysis of all microbial cells, (ii) provide sufficient genomic material (Terrat *et al.*, 2012), and (iii) efficiently remove plant-derived contaminants, particularly PCR-inhibiting phytochemicals (e.g. polysaccharides, polyphenolic compounds, secondary metabolites, etc.) and enzymes (e.g. DNases, proteinases) (Wilson, 1997; Piest *et al.*, 2001; Demeke and Jenkins, 2010). PCR-inhibition occurs when the template DNA is precipitated, degraded, denatured or bound to complex compounds and thus inaccessible to PCR-enzymes. Inhibiting compounds could also bind to the

polymerase and reduce its activity, or plant-derived proteinases could degrade PCR-enzymes (Demeke and Jenkins, 2010).

At present, there is a wide range of plant-tissue DNA extraction protocols available (Table 3.1). Some of these are established conventional laboratory DNA extraction protocols, which utilise detergents (e.g. CTAB and SDS) to lyse cells and liberate genomic material (Demeke and Jenkins, 2010). These protocols can be modified by altering the concentration and/or composition of active ingredients, technical procedures (e.g. mechanical lysis, incubation periods, etc.) and purification processes in order to adapt them to the plant-tissue matrix from which DNA is extracted (Miller *et al.*, 1999; Sharma *et al.*, 2007; Chen *et al.*, 2009). Alternatively, commercial kits have been designed to process specific sample matrices (e.g. soil, fecal, plant and water samples) (Table 3.1).
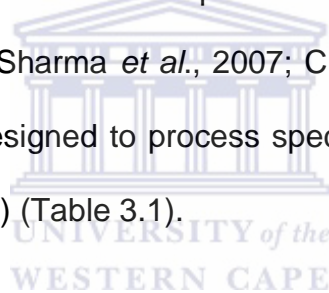
The evaluation of the different extraction protocols for extraction of DNA from plant material has become standard practise. Such studies have been conducted on different plant tissue matrices including lignin-rich woody plants tissues (Green *et al.*, 1999); leaves (Kit and Chandran, 2010); polysaccharide- and phenolic-rich plant tissues (Porebski *et al.*, 1997; Mornkham *et al.*, 2012); flowering organs (Khanuja *et al.*, 1999) as well as seeds (Demeke and Ratnayaka, 2009; Chen *et al.*, 2009). However, the majority of these studies evaluate protocols based on the yield and purity of plant DNA retrieved, as well as its usability in plant genetics studies (Demeke and Jenkins, 2010). Very few, if any, evaluate the effect of extraction protocols on the diversity of plant-associated endophytic bacteria, despite the

number of diversity studies on endophytic bacteria using PCR-based tools such as DGGE (Hardoim *et al.*, 2012; Ramond *et al.*, 2013), t-RFLP (Sessitsch *et al.*, 2012; Ding *et al.*, 2013) and next generation sequencing (Gottel *et al.*, 2011; Lucero *et al.*, 2011; İnceoğlu *et al.*, 2011; Molina *et al.*, 2012).

.

Table 3.1. Partial list of commonly used plant DNA extraction methods and commercial kits (Demeke and Jenkins, 2010).

Extraction protocol	Description	Examples of Plants	Reference
CTAB lysis and purification	Effective for a wide range of matrices. Complexes out polysaccharides and proteins. Used with phenol/chloroform: isoamyl alcohol. Takes a long time	A wide range of plants and matrices	Terry *et al.*, 2002; Doyle and Doyle, 1987; IOS, 2005
DNeasy (mini and maxi-kits) (Qiagen)	Uses silica-gel-membrane and simple spin procedures to isolate DNA.	Corn, soybean	Corbisier *et al.*, 2007; Cankar *et al.*, 2006; Holden *et al*, 2003
Epicentre MasterPure™ Complete (EPICENTRE Biotechnologies)	Utilizes rapid salt precipitation protocol to remove contaminating macromolecules	Soybean, maize	Bernado *et al.*, 2007
Fast ID genomic DNA extraction kit/Fast ID food DNA extraction kit (Genetic ID)	Uses genomic lyse and bind buffers as well as DNA binding columns	Soybean, maize, rice, processed food and feed samples	Chhalliyil *et al.*, 2008
GenElute™ plant genomic DNA miniprep kit (Sigma)	Uses silica binding and elution in a spin column format	Maize flour	Rizzi *et al.*, 2003
GM quicker (Nippon Gene; bioMerieux)	Uses silica spin column and anionic detergent	Soybean, maize	Corbisier *et al.*, 2007; Minegishi *et al*; 2008
PVP method	Includes thermal lysis in the presence of SDS and high EDTA concentration followed by removal of contaminants such as polyphenolic compounds and polysaccharides.	Used for various matrices	IOS, 2005
QIAamp DNA stool kit (Qiagen)	Silica-membrane-based purification system. Suitable for PCR-inhibitor-rich substances and for highly processed foodstuffs	Soybean, maize	Peano *et al.*, 2004; Tengel *et al.*, 2001
QIAEX II gel extraction kit (Qiagen)	Based on selective adsorption of nucleic acids onto QIAEX II silica-gel particles in the presence of chaotropic salt. Used for DNA purification. Lipase treatment may be required for fat-rich matrices	Soybean, canola	Terry *et al.*, 2002; Demeke and Ratnayaka, 2008
SDS	Improves lysis of cells. Widely used for DNA extraction from seeds. SDS is also used in combination with phenol/chloroform	Used for a wide range of plants	IOS, 2005; Delaporta *et al.*, 1983
UltraClean plant DNA kit (Mo-Bio)	Cell lysis achieved with bead grinding; silica binding; spin column format.	Potato	Smith *et al.*, 2005

An important consideration in metagenomic studies of plant associated bacterial communities, is that genomic material retrieved should be representative of microbial communities found in the habitat being studied (Terrat *et al.*, 2012). Previously, protocols were compared on their effect on microbial diversities found in environments such as soil (Sessitch *et al.*, 2002; de Lipthay *et al.*, 2004; Hirsh *et al.*, 2010), clinical samples (Fredericks *et al.*, 2005; Willner *et al.*, 2012) and stool samples (Salonen *et al.*, 2010), but reports of similar comparisons on plant tissues are lacking.

This study is, therefore, the first to report on the effect of DNA extraction protocols on endophytic bacterial communities associated with sorghum, pearl millet and groundnut tissues. The aim is to select a protocol that introduces the least bias in bacterial community diversity analyses, for use in subsequent studies of endophytic bacterial communities associated with these crops. Two classical protocols (CPs) and five commercial kits (CKs) are compared for the yield of DNA obtained from low (0.1g) and high quantity (0.3g) plant tissue, and its quality for use in PCR-based analyses. In addition, the diversity of retrieved endophytic bacterial communities is assessed using t-RFLP. The two classical protocols were either SDS-based (Zhou *et al.*, 1996) or CTAB-based (Murray and Thompson, 1980). Three of the commercial kits (MoBio PowerPlant Pro® DNA Isolation Kit, Qiagen DNeasy^R Plant Mini Kit and Fermentas GeneJET Plant Genomic DNA Purification Kit) were designed for plant DNA extraction (CK_p) and two (MoBio PowerSoil^{TM} DNA Purification Kit and MoBio UltraClean® Soil DNA Isolation Kit) for soil DNA extraction (CK_s).

3.2. Results

3.2.1. Effects of DNA extraction methods on yield and quality of plant metagenomic DNA

3.2.1.1. Effects of extraction protocols on DNA yield

When 0.1g starting plant material was used, CPs extracted metagenomic DNA (mDNA) of high molecular weight (approximately 15-20kbp) from the different plant tissues (Figure 3.1). DNA retrieved with CKs was of lower molecular weight (less than 15kbp). Based on electrophoresis visualisation of extracted DNA, more mDNA was retrieved with CPs than with CKs. CK_ss provided the least amount of DNA, which was almost undetectable on the agarose gel, for the same amount of starting plant tissue material. Also, more DNA was extracted from groundnut tissues than from sorghum or pearl millet tissues. DNA retrieval patterns at low (0.1g) and high (0.3g) (data not shown) starting plant material were similar for all protocols.

DNA yield was further quantified via Nanodrop measurements. When low (0.1g) starting plant material was used, yield of mDNA extracted with classical protocols (CPmDNA) was significantly higher than yield obtained with commercial kits (CKmDNA) (p<0.001) across all tissue types (Figure 3.2A). The SDS-based protocol, in particular, provided the highest DNA yields [48.08 (±8.59)ng.mg^{-1} to 490.78 (±80.03)ng.mg^{-1}] from all tissues. However, DNA yields obtained with this protocol were also the least reproducible. The CTAB-based protocol retrieved significantly

less DNA than the SDS-based protocol (p<0.001) [13.74 ($\pm$1.1)ng.mg^{-1} to 91.30 ($\pm$7.0)nm.mg^{-1}].
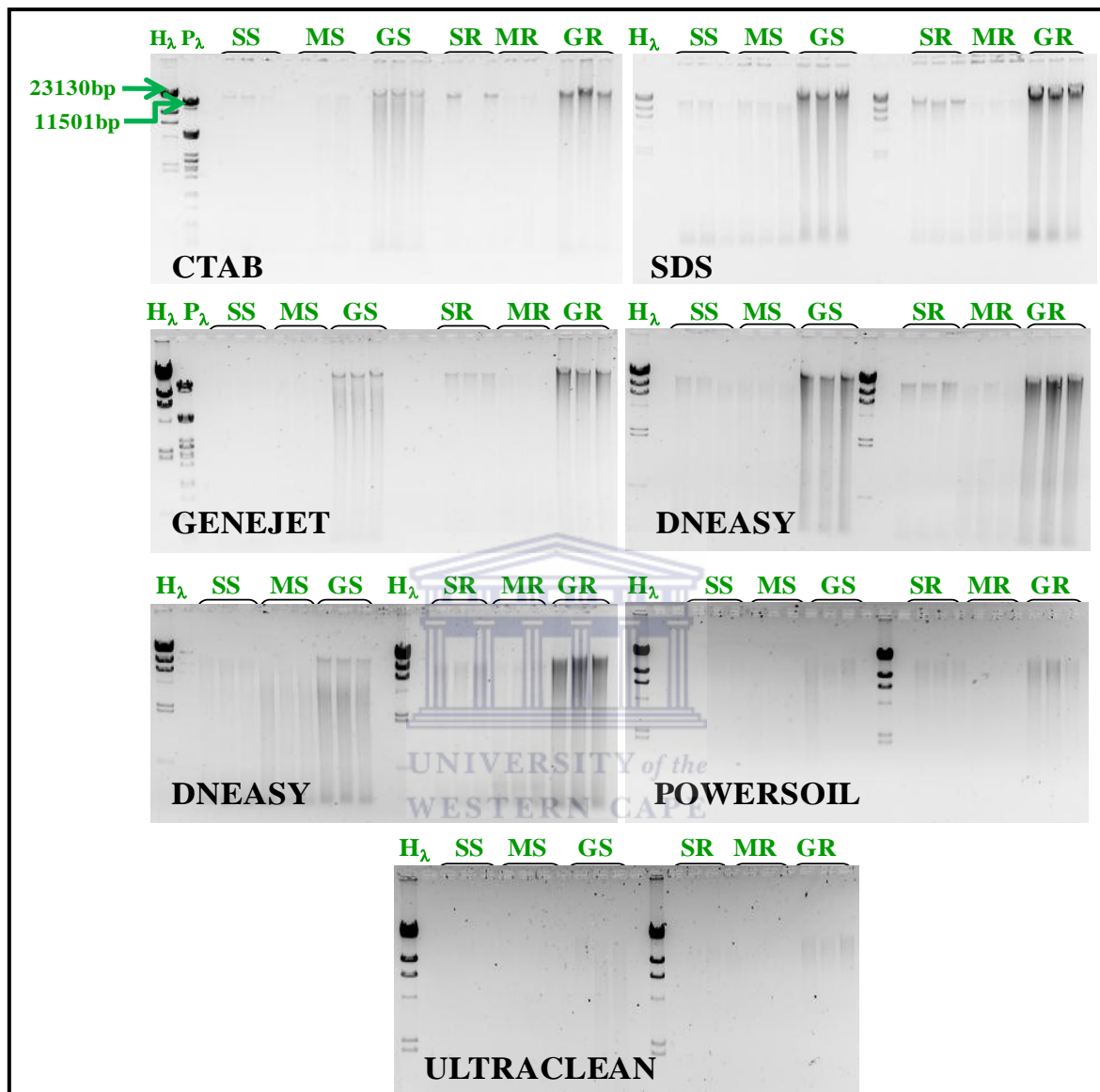
Figure 3.1. Visualisation of mDNA on 0.8% agarose gels after electrophoresis. DNA was extracted using seven extraction protocols from 0.1g sorghum stem (SS), sorghum root (SR), pearl millet stem (MS), pearl millet root (MR), groundnut stem (GS) and groundnut root (GR). DNA molecular size is determined by comparison to molecular markers, lambda DNA digested with *Hind*III (H$_\lambda$) and *Pst*I (P$_\lambda$) restriction enzymes. Extractions were performed in triplicate.

When comparing CKs, CK_ps provided significantly higher DNA yields across all plant tissues than (CK_ss) ($p<0.001$) (Figure 3.2A$_i$). Sorghum and pearl millet tissue mDNA yields were not significantly different ($p>0.05$) across all CK_ps, [3.73 (±0.1) to 6.19 (±0.6)ng.mg^{-1}]. Yield of groundnut DNA extracted with the PowerPlant kit [28.59 (±1.4)ng.mg^{-1}] was significantly lower when compared to yields retrieved with the GeneJet and DNeasy kits ($p<0.05$). DNA yields obtained with the two CK_ss were not significantly different across all tissue types ($p>0.05$), and were also the lowest at less than 28.60 (±1.44)ng.mg^{-1}.

DNA yield comparisons were repeated using 0.3g starting plant material. Only protocols that provided the sufficient yields previously, i.e. CPs and CK_ps, were used. At a higher plant material quantity, CPmDNA yield was still significantly higher than CKmDNA yield ($p<0.001$). There was no significant change in sorghum or pearl millet mDNA yield when the SDS protocol was used. Groundnut mDNA yield was significantly reduced to 99.48 (±7.9)ng.mg^{-1} and 93.95 (±9.5)ng.mg^{-1} for the stem and root tissues, respectively. No significant change in yield was observed across all tissues processed with the CTAB protocol or the K_ps, when starting plant material was increased from 0.1g to 0.3g.

All pair-wise comparisons between plant tissues across all protocols at 0.1g or 0.3g starting plant material showed that there was no significant difference in mDNA yields between tissues of the two monocotyledonous plants, sorghum and pearl millet (Holm-Sidak test, all p-values > 0.05). Contrastingly, dicotyledonous groundnut

mDNA yields were significantly higher than yields obtained from the monocots (all p-values < 0.001).
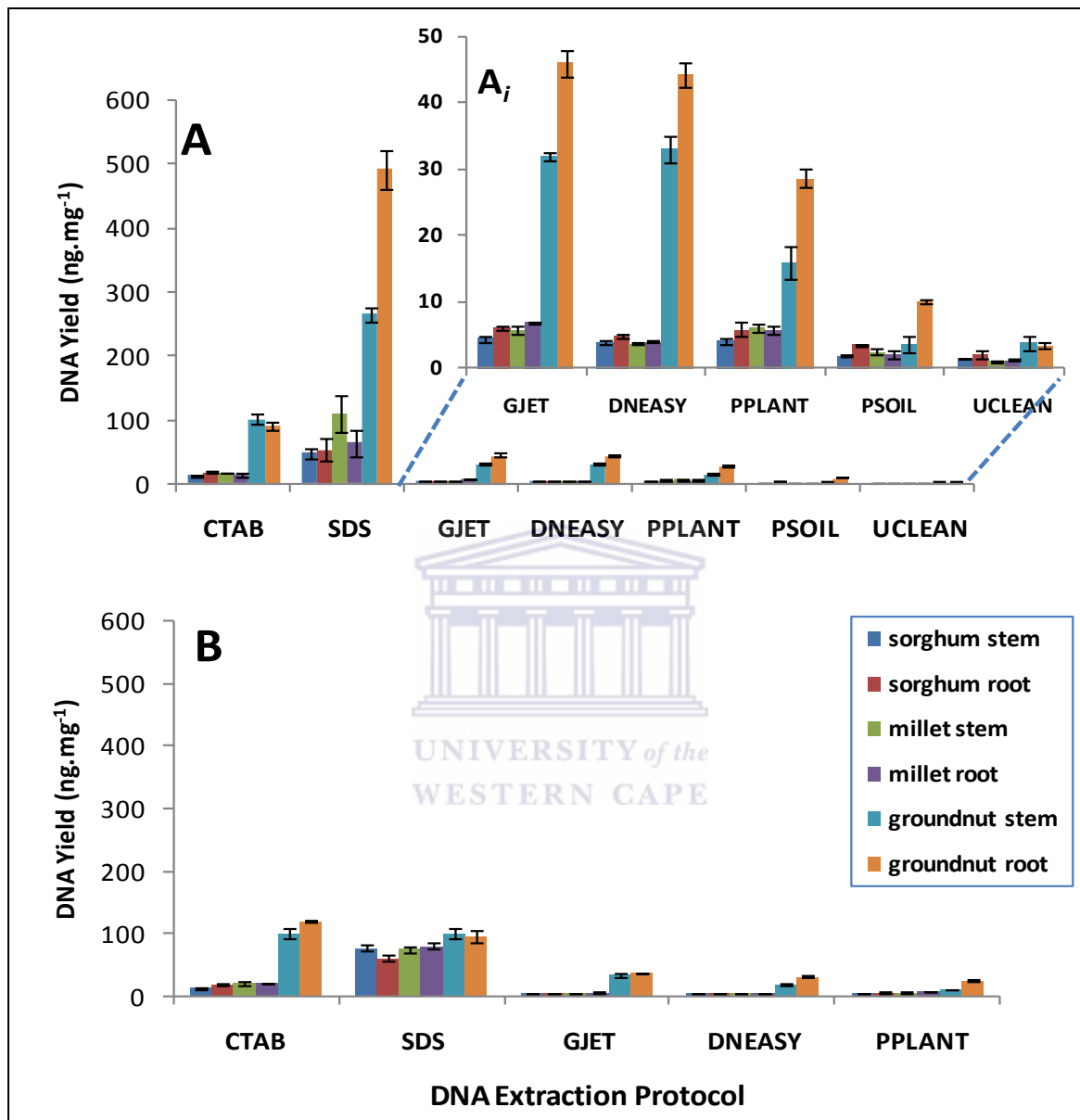
Figure 3.2. Metagenomic DNA yield from 0.1g (A) and 0.3g (B) sorghum, pearl millet and groundnut tissues using classical protocols and commercial kits. In both tests, CPmDNA yield is significantly higher than CKmDNA when compared at 95% significance level (ANOVA, p<0.001).

3.2.1.2. Effects of extraction protocols on DNA purity and quality

Metagenomic DNA was considered pure when the ratio of absorbance at wavelength 260nm and 280nm was between 1.7 and 1.9. Figure 3.3 compares the purity estimates of mDNA samples based on this ratio.
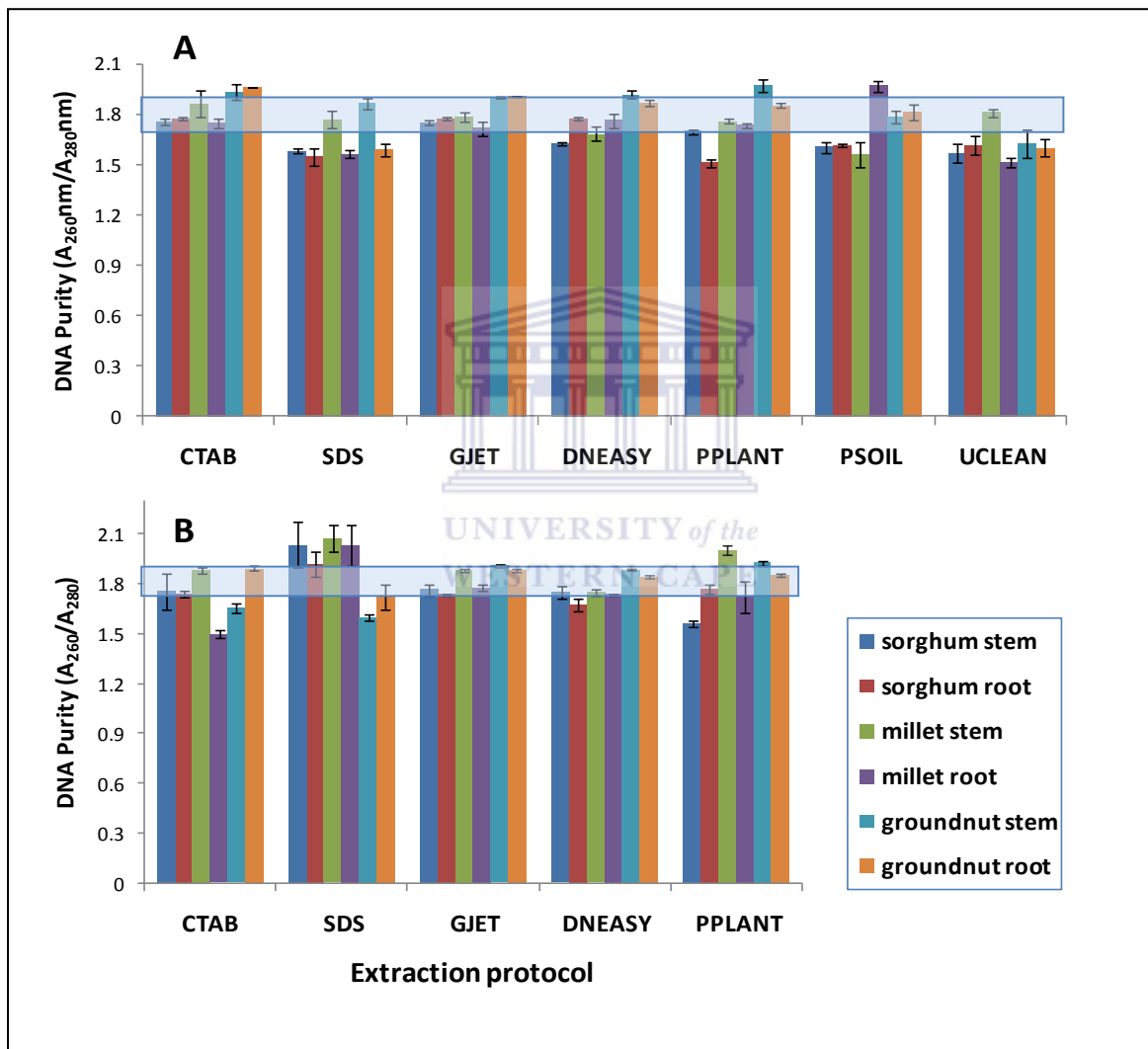
Figure 3.3. Purity estimation for metagenomic DNA extracted from 0.1g (A) and 0.3g (B) sorghum, pearl millet and groundnut tissues with classical and kit protocols. DNA is considered pure when nanodrop-calculated A_{260}/A_{280} ratio is between 1.7 and 1.9 (range shown as horizontal blue strip).

CPs were not consistent in extracting mDNA of high purity when 0.1g starting plant material was used (Figure 3.3A). Sorghum and pearl millet mDNA extracted with the CTAB protocol was pure; however, groundnut root and stem mDNA extracted using the same protocol was not pure. Of all samples processed with the SDS protocol, only pearl millet stem mDNA samples were pure.

DNA purity varied between CKs brands. The GeneJet kit performed best in this regard, as it extracted pure mDNA from all plant tissues. When the PowerPlant kit was used, only sorghum root and groundnut stem mDNA samples were impure, whereas the DNeasy kit only retrieved pure mDNA from the three root samples. The PowerSoil kit extracted pure mDNA from groundnut stem and root only, and the UltraClean Kit from millet stem only.
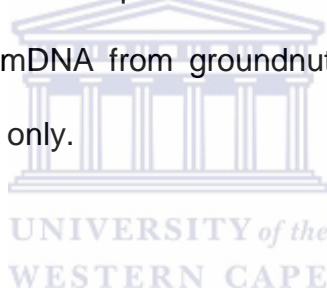
When starting plant tissue material was increased to 0.3g, the mDNA quality and reproducibility of results deteriorated with the SDS protocol (Figure 3.3B). Pure mDNA was retrieved from groundnut root only. When the CTAB protocol was used, only millet root and groundnut stem mDNA samples were impure. Purity of mDNA extracted with the GeneJet kit was not affected by an increase in starting plant material as all samples were sufficiently pure. When the DNeasy protocol was used, only sorghum root mDNA was found to be impure. The PowerPlant kit could only retrieve pure mDNA from root samples of all three plants.

Nanodrop analysis of DNA is only a preliminary step in determination of DNA quality. It indicates the possibility of presence or absence of impurities in the DNA. However, it has been shown that slight variations in the buffer pH and ionic strength can affect the light absorbance by DNA, and thus alter the A_{260}/A_{280} measurements (Wilfinger *et al.*, 1997). These changes can be easily introduced through technical handling of DNA solutions. Also, process-introduced DNA contaminants that produce similar A_{260} and A_{280} absorbance profiles to DNA (e.g. guanidine hydrochloride) could be undetected when using this technique (ThermoScientific, T042). Manufacturers of commercial kits (e.g. Qiagen) include Guanidine-HCl as a protein digesting ingredient in purification buffers (Terry *et al.*, 2002). In light of these concerns, purity was also indirectly tested by using different mDNA amounts (1ng, 5ng and 10ng) as template in PCR reactions. These tests assist in determining whether the mDNA contains PCR-inhibiting compounds.
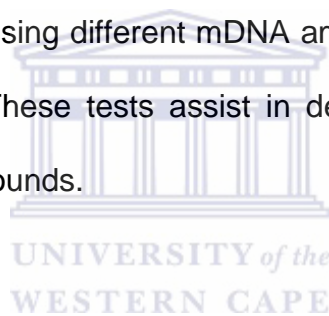
When low starting plant material (0.1g) was used, PCR inhibitions were most frequently observed when CPmDNA was used as template in PCR amplification of the bacterial 16S rRNA gene (for example, Figure 3.4). PCR inhibition was not correlated to amount of template mDNA used. For example, inhibition occurred when sorghum stem mDNA samples 1 (10ng, 1ng), 2 (10ng, 1ng) and 3 (1ng) extracted with CTAB protocol were used. Similar observations were made when sorghum root, millet root or groundnut tissue mDNA were used; however, amplification was successful with all pearl millet stem samples. When the SDS protocol was used, PCR-inhibition was observed for one sorghum stem mDNA (sample 3, 5ng) and one sorghum root mDNA (sample 1, 10ng). With pearl millet samples, inhibition was observed for stem samples 1 (1ng) and 2 (10ng) as well as pearl millet root sample 2

(1ng). Groundnut mDNA inhibitions were observed for stem samples 2 (5ng) and 3 (1ng), and for root samples 2 (10ng, 5ng, 1ng) and 3 (10ng).
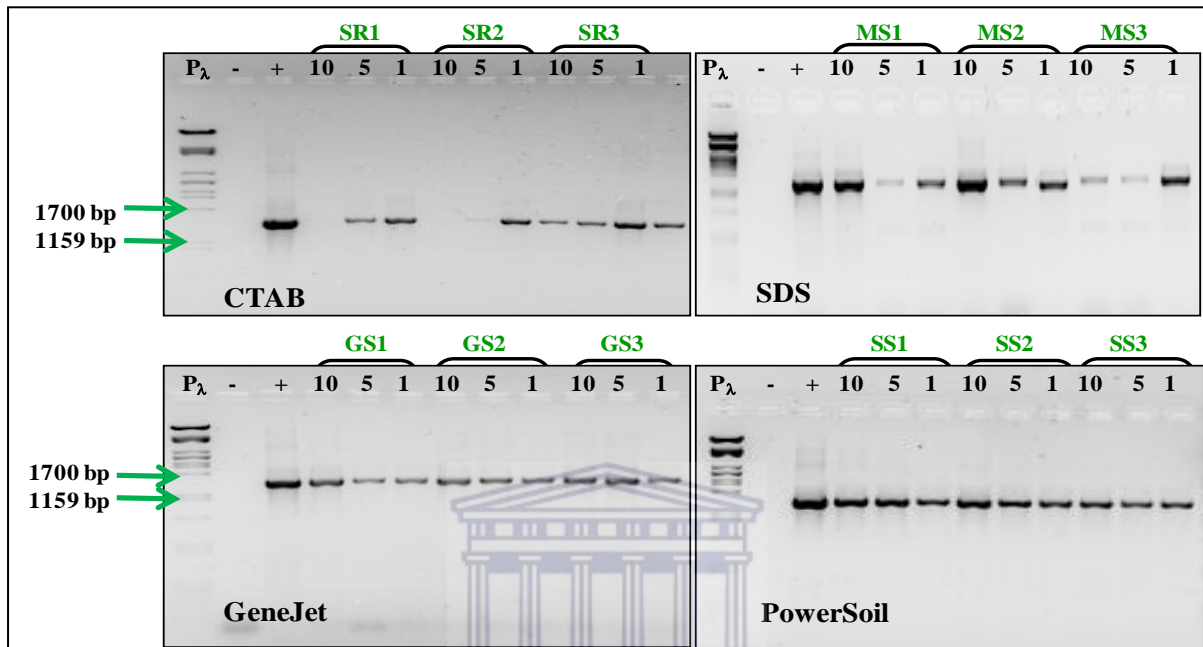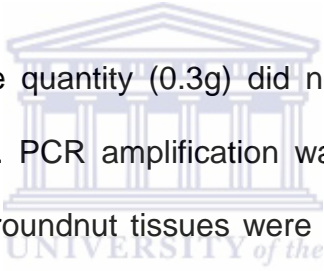
Figure 3.4. 16S rRNA amplification from metagenomic DNA extracted from 0.1g plant tissue. SR = sorghum root, MS = pearl millet stem, GS = groundnut stem and SS = sorghum stem. Amplicon size is determined by comparison to *Pst*I lambda DNA (P_λ) molecular marker. Double-distilled molecular grade water was used as negative control and *Geobacillus* sp. genomic DNA (5ng) was used as positive control (+). All amplifications were conducted in triplicate.

Less PCR-inhibitions were observed when CKmDNA was used. PCR amplification was successful with all sorghum and groundnut mDNA samples extracted with the GeneJet kit. One PCR inhibition incident was observed with pearl millet stem sample 3 (5ng). With mDNA extracted with the DNeasy kit, only two inhibitions were

observed, pearl millet root sample 1 (1ng) and 2 (5ng). No inhibition was observed when mDNA (all tissues) extracted with the PowerPlant kit was used. PCR amplification was successful with all sorghum mDNA samples extracted with the PowerSoil kit, and only one inhibition was observed with pearl millet mDNA (1ng, root sample 3). Two inhibitions were observed when groundnut mDNA was used (1ng, stem sample 3 and 5ng, root sample 3). No inhibitions were observed when sorghum and pearl millet mDNA samples retrieved with the UltraClean kit were used. Inhibitions were observed for three groundnut mDNA samples (1ng of stem samples 1 and 3; 1ng root sample 1).

Increased starting plant tissue quantity (0.3g) did not affect the quality of mDNA extracted with CTAB protocol. PCR amplification was successful when all mDNA samples from sorghum and groundnut tissues were used. PCR inhibition was only observed with pearl millet mDNA samples 1(10ng), 2(1ng) and 3(1ng). Amplification was also achieved with all pearl millet (root and stem) and sorghum stem mDNA samples retrieved with the SDS protocol. Inhibition was observed with one pearl millet root mDNA sample (1ng) and all groundnut mDNA samples. These results, together with purity estimations in Figure 3.2B, strongly suggest that mDNA extracted with the SDS protocol from 0.3g groundnut tissue contains impurities that hinder PCR reactions. Groundnut mDNA was then further purified using PVPP spin columns and the PCR test was repeated. After purification stem and root mDNA purities were improved, with A_{260}/A_{280} ratios of 1.69 (±0.04) and 1.81 (±0.05), respectively. Amplification of the bacterial 16S rRNA was also improved as fewer inhibitions were observed when stem [sample 2 (5ng), sample 3 (1ng)] and root [samples 1-3 (1ng)] mDNA were used. However, the additional purification step

reduced the stem and root mDNA yields to 62.27 ($\pm$4.36)ng.mg^{-1} and 67.88 ($\pm$7.30)ng.mg^{-1}, respectively.

The data presented has shown that DNA extraction protocols affect yield and quality of mDNA extracted from plants. Also, estimation of DNA purity by measurement of spectrophotometric absorbance alone is not sufficient in determining the quality of mDNA used for PCR applications. Based on these preliminary findings, the CTAB protocol and CK$_p$s provided mDNA of sufficient yield and purity, from all tested plant tissues, for use in downstream PCR-based applications. The SDS protocol also provided high-yield mDNA, but only the sorghum and pearl millet mDNA was of good quality for PCR applications. The quality of groundnut mDNA extracted with this protocol was poor for use in PCR reactions. The yield of mDNA extracted with CK$_s$s is insufficient for use in metagenomic studies, even though the quality was adequate for PCR applications. However, bulking up on the starting material in future could solve this.

3.2.2. Effects of extraction protocols on endophytic bacterial community diversity

T-RFLP was used to compare the diversities of endophytic bacteria associated with sorghum and groundnut tissues. Closely related and structurally similar monocots, sorghum and pearl millet plants (both belonging to the Poaceae family), provided similar results in previous analyses (section 3.2.1), thus indicating that their mDNA samples were similarly affected by different DNA extraction protocols. Presumably, these protocols will also affect the endophytic diversity in a similar manner. Pearl

millet tissues were thus excluded in the current tests, and sorghum communities were compared with endophytic communities of a distant plant, groundnut (family Fabaceae). One CK_p, (PowerPlant) and one CK_s (PowerSoil) were also excluded.

3.2.2.1. Diversity of retrieved endophytic communities

Species richness (S), Shannon index (H') and the Simpson index (1-λ) were used to measure the diversity of communities retrieved from sorghum and groundnut tissues (Clarke & Warwick, 2001). S is a direct count of operational taxonomic units (OTUs) observed. H' measures the proportion of all OTUs in the whole community, and it is calculated as H' = -$\sum_i p_i$ log (p_i), whereby p_i is the proportion of the total count arising from the ith OTU. (1-λ) measures community evenness (or equitability), which quantifies how evenly distributed OTUs are within a community. It is calculated as 1-λ= 1-[$\sum_i N_i$(N-1)} / [N(N-1)], where N_i is the number of OTUs that belong to species i.

When 0.1g starting plant material was used, the number of OTUs (S) retrieved with CPs [7 (±1) to 10 (±2)] from sorghum and groundnut tissues was significantly greater than OTUs retrieved with CKs [2 (±1) to 7 (±3)] (p-values < 0.05) (Figure 3.5, A1). There was no significant difference in S or H' between communities retrieved with the CTAB- and the SDS-based protocols from individual plant tissues (Figure 3.5, A1 and B1). Plant tissue type did not have a significant effect on the number of OTUs retrieved (p=0.066).
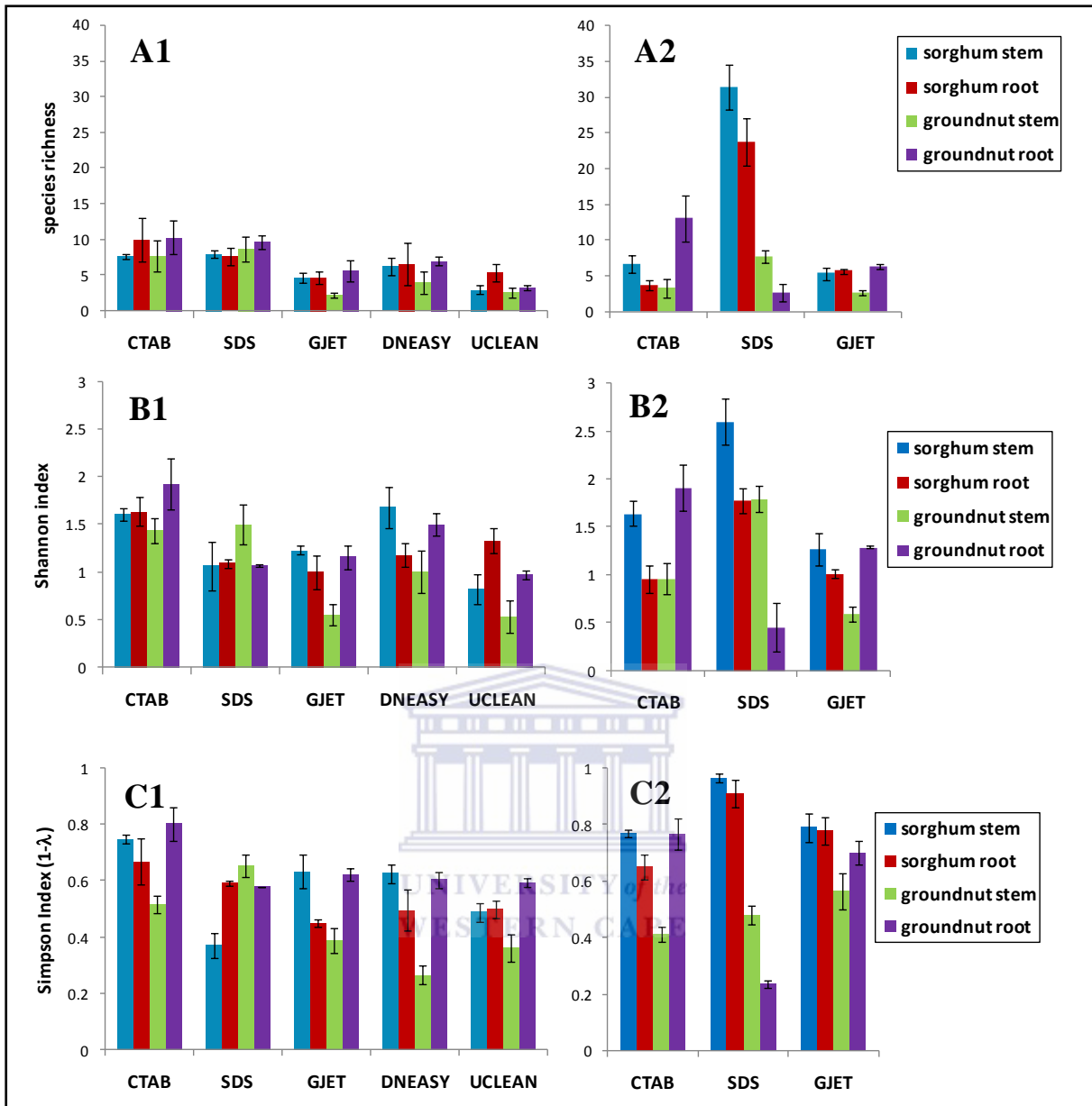
Figure 3.5. Calculated diversity indices, species richness (S), Shannon Index (H')
and Simpson Index (1-λ) for endophytic communities retrieved from 0.1g (A1, B1,
C1) and 0.3g (A2, B2, C2) root and stem tissues of sorghum and groundnut.

However, sorghum communities retrieved with the CTAB protocol were more evenly
distributed [0.67 (±0.08) ≤ (1-λ) ≤ 0.80 (±0.06)] than those retrieved with the SDS
protocol [0.37 (±0.04) ≤ (1-λ) ≤ 0.65 (±0.04)] (Figure 3.5, C1). This implies that the

SDS protocol is biased towards certain bacterial groups found in these tissues. Groundnut stem community evenness values were midrange with both the CTAB protocol 0.52 (±0.03) and the SDS protocol 0.65 (±0.03).

There was no significant difference in the number of OTUs retrieved with the CKs (all p-values > 0.05) (Figure 3.5, A1). The DNeasy kit was able to access more unique OTUs from all tissues [$1.008(\pm0.22) \leq H' \leq 1.68(\pm0.22)$] compared to the GeneJet Kit [$0.552(\pm0.11) \leq H' \leq 1.232(\pm0.05)$] and the UltraClean kit [$0.539(\pm0.17) \leq H' \leq 1.334(\pm0.13)$]. All evenness measurements for communities retrieved with CKs were low to mid-range [$0.26(\pm0.03)$ to $0.63(\pm0.05)$], indicating greater dominance by certain phylotypes (Figure 3.5, C1).
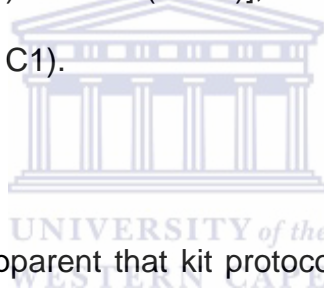
From these findings, it was apparent that kit protocols compromise the diversity of endophytic communities more than classical protocols. When starting plant material was increased, the classical protocols were then compared only to the GeneJet kit as it was shown to have consistent relative efficiency (compared to other kits) based on DNA yield and purity tests, as well as current diversity analyses.

Significant decreases in S and H' values of sorghum root and groundnut stem communities (P-values < 0.05) retrieved with the CTAB protocol were observed at a high starting plant quantity (0.3g) (Figure 3.5, A2 and B2). There was no significant change in S and H' for sorghum stem and groundnut root communities. No significant change in evenness across all tissues (p-values > 0.05) was observed

(Figure 3.5, C2). S, H' and evenness of sorghum stem and root endophytic communities were significantly increased when the SDS protocol was used. Species richness was 31 (±3) and 23 (±3) OTUs, for sorghum stem and roots respectively, and H' values were 2.60 (±0.24) and 1.77 (±0.13). Evenness was approaching maximum equitability at 0.96 (±0.01) and 0.91 (±0.05) for the stem and the root respectively. This suggests that the SDS protocol was able to access more bacteria in sorghum tissues than it did when 0.1g plant tissue was used.   Contrastingly, the diversity of groundnut communities retrieved with the SDS protocol was negatively affected by an increase in starting plant material. There was no significant change in groundnut stem species richness and H' value, whilst root community species richness was significantly reduced to 3 (±2) OTUs and H' to 0.46 (±0.25). Community evenness decreased to between 0.24 (±0.01) and 0.48 (±0.03). There was no significant change in S and H' values of communities retrieved with the GeneJet kit across all plant tissues. However, community evenness was increased by at least an order of magnitude across all tissues. This implies that an increase in starting plant tissue only marginally increased the diversity of communities accessed with this protocol.

3.2.2.2.   Ordination of bacterial communities using non-metric multi-dimensional scaling (nm-MDS) plots

A non-metric MDS plot allows ordination of community samples according to their dissimilarity measurements, i.e., the closer samples are on the nm-MDS plot, the more similar they are to each other (Clarke and Warwick, 2001). The nm-MDS stress value measures the goodness-of-fit of the non-parametric regression line from which

sample dissimilarities are calculated. Therefore, it measures reliability of the ordination plot (Clarke and Warwick, 2001).

Ordination of endophytic bacterial communities retrieved from 0.1g starting plant material revealed two major clusters (Figure 3.6). The nm-MDS stress value of 0.07 indicates good ordination, with minimal prospects of misinterpretation (Clarke and Warwick, 2001). The extraction protocols used had a significant effect on the structure of endophytic bacterial communities retrieved (ANOSIM, R = 0.441, p < 0.001). One cluster indicated high similarity between all communities retrieved with the commercial kits and the CTAB-protocol, as well as sorghum root communities retrieved with the SDS-protocol (Figure 3.6). Statistical analysis of community structures confirmed that communities retrieved with CKs were not significantly different (-0.019 ≤ R ≤ 0.13). This means that these kits were able to access similar bacterial phylotypes from the different plant tissues. However, statistical comparisons showed that communities retrieved with the CPs were significantly different from communities retrieved with the CKs (0.375 ≤ R ≤ 0.815; all p-values ≤ 0.008). The differences observed between CTAB and CK communities could be due to the additional phylotypes that the CTAB protocol was able to access as previously shown by the calculated diversity measurements.
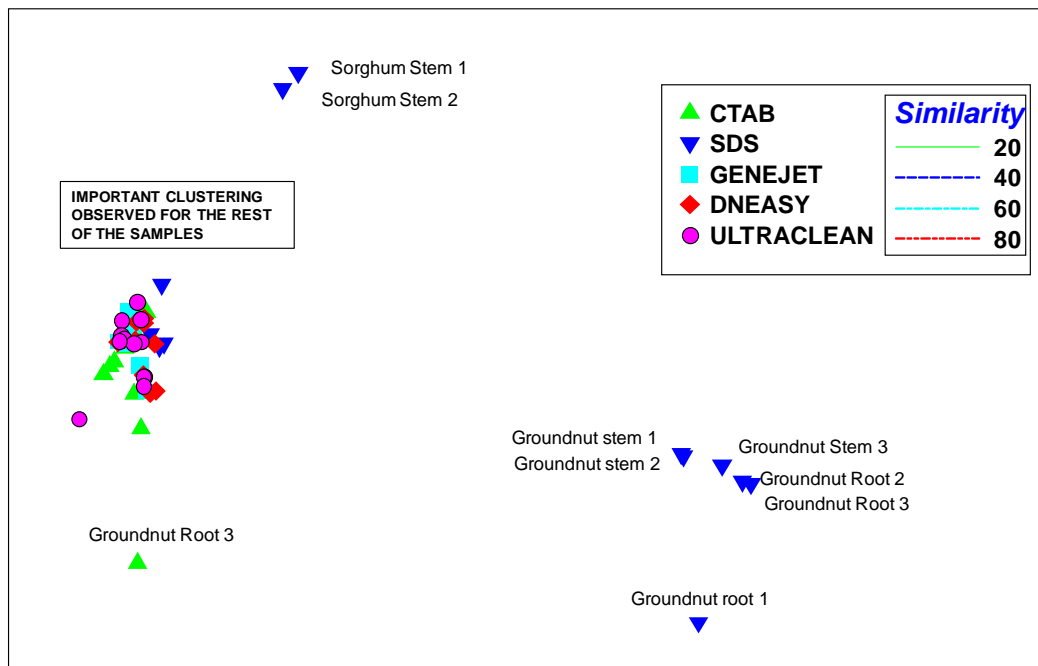
Figure 3.6. Two-dimensional non-metric Multi-Dimensional Scaling (nm-MDS) plot of Bray-Curtis similarity (presence-absence transformation) of communities retrieved using different DNA extraction protocols from low starting plant tissue material (0.1g). Stress = 0.07.
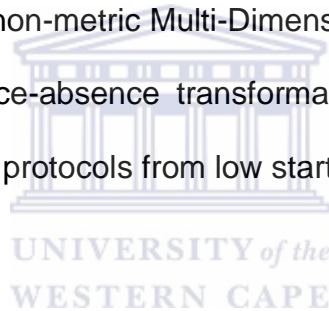
The second cluster consisted of groundnut tissue communities retrieved with the SDS-protocol. Indeed, statistical comparisons confirmed that communities retrieved with the SDS-protocol were significantly different from communities retrieved with the CTAB protocol (R=0.661, p<0.001) or commercial kits (R>0.648, p<0.001). Plant tissue type also had a significant effect on the endophytic bacterial community structure (R = 0.625, p < 0.001).

When starting plant material was increased to 0.3g, endophytic communities were shaped by extraction protocol (Global R = 0.522, p < 0.001) used as well as plant

tissue type (Global R = 0.616, p < 0.01) from which they were derived (Figure 3.7). The nm-MDS value of 1.6 is indicative of some degree of scatter; however, at this stress-value the MDS-plot is still reliable in discerning community structure patterns, particularly when supported by statistical comparisons (Clarke and Warwick, 2001).
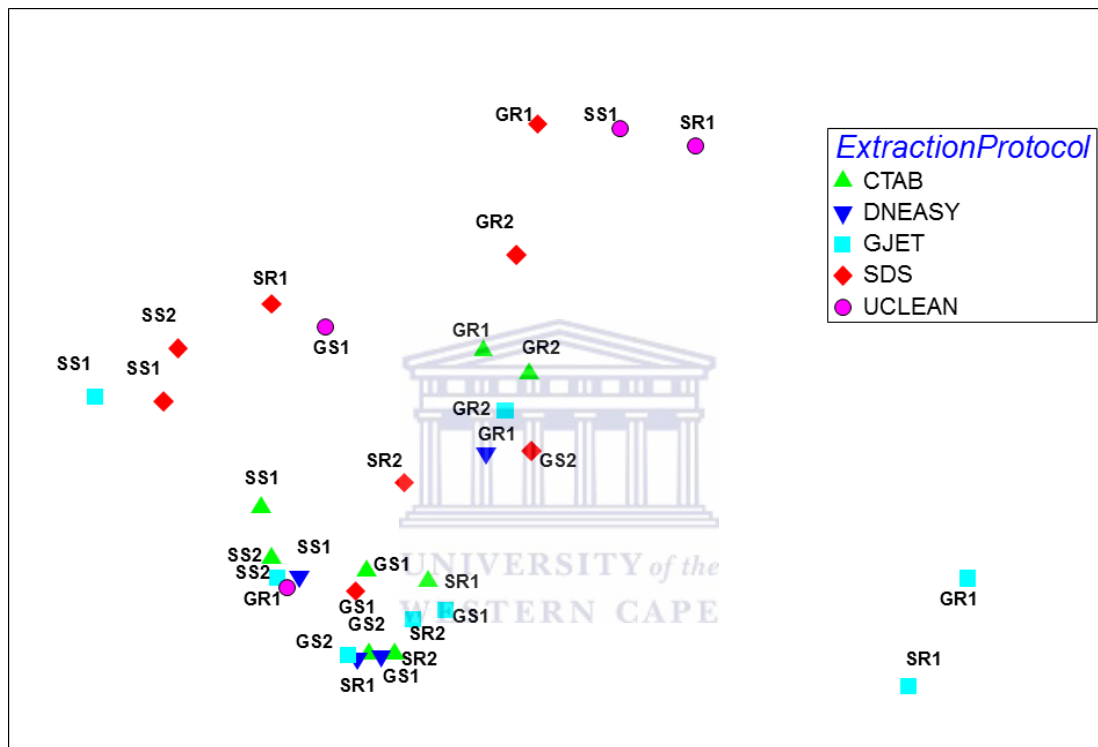
Figure 3.7. Two-dimensional nm-MDS plot of Bray-Curtis similarity (presence-absence transformation) of communities retrieved using different DNA extraction protocols from high starting plant tissue material (0.3g). Stress = 0.16.

Results obtained with the SDS protocol were most reproducible as shown by clustering of replicated samples (e.g., sorghum stem communities SS1, SS2, SS3). Also, with the SDS protocol, samples from different plant tissues were distant from each other on the plot. This indicates that the SDS-protocol was able to retrieve

endophytic communities of different structures from the root and stem tissues of sorghum and groundnut plants. Communities retrieved with the SDS-protocol were significantly different from communities retrieved with the CTAB protocol (R=0.778, p<0.01) or the GeneJet Kit (R=0.556, p<0.01). In contrast, there was no significant difference in communities retrieved with the CTAB and the GeneJet kit (R=0.278, p=0.01). However, reproducibility of results obtained with these two protocols was poor compared to the SDS-protocol.

Inspection of the OTU matrix revealed that none of the protocols were able to access the total OTUs observed. This suggests that none of the mDNA samples retrieved with each individual protocol was representative of indigenous endophytic communities in their entirety, although the degree of bias varied between protocols as shown by diversity measurements (Figure 3.5). An attempt was made to identify bacteria retrieved by each protocol from specific plant tissues by *in silico* digestion, in order to determine groups that are most favoured by different protocols. However, this was unsuccessful because all individual t-RFs matched a broad range bacterial phylotypes, thus indicating the inability of t-RFLP to resolve taxonomy at a fine scale.

From these tests, it is apparent that the efficiency of DNA extraction protocols in metagenomic studies of plant-associated endophytic bacteria is affected by the plant species as well as the size of the starting plant tissue material. In summary, CPs are efficient in providing high-yield mDNA and CKs provide mDNA of superior PCR quality. However, mDNA extracted with CKs under-represent community diversities. Metagenomic DNA extracted with the SDS protocol best represents diversities of

sorghum-associated indigenous bacterial communities; particularly when mDNA is extracted from higher plant tissue quantity (0.3g). Therefore, this protocol is recommended for culture-independent ecological studies of these endophytes. Groundnut community diversities were best represented with mDNA extracted with the CTAB protocol when lower plant tissue quantity was used (0.1g).

3.3. Discussion

DNA extraction protocols constitute one of the most critical components of molecular studies of microbial communities. This is because accurate estimation of community profiles and diversity depends on the ability of the extraction protocol to retrieve mDNA that is usable in PCR-based analyses and is representative of indigenous microbial populations (Terrat *et al.*, 2012). This is particularly important in studies of plant-associated endophytic microorganisms, whereby the extracted DNA is susceptible to plant-derived contaminants (Demeke and Jenkins, 2010). The accessibility of microbial DNA is further compromised by plant tissue structures, whereby upon extraction, it is mixed with the plant's own genomic material. In addition, plant tissue matrices differ by plant species and plant organs.

It is therefore important for microbial ecologists to select a procedure that introduces the least biases when studying endophytic bacterial communities. One advantage of classical protocols is that they are amenable to modifications and can therefore be optimised to process a wider range of plant matrices (Sharma *et al.*, 2007; Chen *et al.*, 2009). These processes are, however, time-consuming and use toxic reagents

such as phenol and chloroform. Also, they have been shown to introduce variability in DNA yield and quality (Salonen *et al.*, 2010). Consequently, commercialised DNA extraction kits were developed to standardise the DNA extraction procedure. Kit protocols are also faster and use non-toxic reagents.

The common thread in all published comparative studies of plant DNA extraction protocols is that they are end-use specific; i.e, factors tested in these studies are selected based on the intended use of the extracted DNA. For example, Drabkova *et al.* (2002) compared seven extraction protocols (3 classical protocols and 4 kits) on the yield of intact DNA from herbarium Juncaceae plant collections of different ages for use in plant phylogenetic studies. Mornkam *et al.*, (2012) evaluated the use of mDNA extracted from seeds and leaves of phenolic- and polysaccharide-rich Jerusalem artichoke plant in PCR-amplification of specified regions on the plant genome. Four classical protocols and one commercial kit were compared on retrieval of mDNA that can be used in the study of a broad range of targeted plant genes. Effects of different DNA purification processes on plant DNA intended for plant genomic studies have also been compared (Demeke and Ratnayaka, 2009). These comparative studies were, however, mostly restricted to quality analysis of DNA intended for plant genomic studies. To our knowledge, there are very few (if any) comparative studies that evaluate plant-derived mDNA for phylogenetic characterization of plant-associated bacteria. This study would therefore be the first to analyse the effects of DNA protocols on the diversity of endophytic bacteria associated with sorghum and groundnut plants.

Previous comparisons between classical protocols and commercial kits strongly suggest that classical protocols are more superior in extracting DNA of high-yield from plant tissues (Drabkova *et al.*, 2002; Niu *et al.*, 2008; Sahu *et al.*, 2012). This concurs with observations made in the current study. DNA yield relies on the efficiency of the cell lysis step, which involves mechanical (e.g. grinding, dead-mill) and chemical (e.g. enzymatic lysis) processes to disrupt cells (Moré *et al.*, 1994). When studying plant associated endophytes, "harsh" lysis techniques such as grinding and bead-mills combined with chemical cell lysing reagents would be more effective in disruption of the plant tissue structure (plant matrix) as well hard cell wall bacteria such as the endospore-forming *Bacillus subtilis* (Moré *et al.*, 1994; Yuan *et al.*, 2012). This is to increase access to more genetic material and a wider diversity of microorganisms (Yuan *et al.*, 2012).

In this study, plant tissues were first subjected to mechanical grinding under liquid nitrogen. Plant DNA extracting kit protocols recommend the use of a bead-mill homogeniser for increased yield of DNA, and its efficiency was shown in previous studies (Miller *et al.*, 1999). However, for the benefit of consistency in comparisons, this tool was not used in this study. Mechanical cell lysis procedures employ rigorous force to disrupt cells, but this can also lead to DNA shearing (Varma *et al.*, 2007). SDS/lysozyme and CTAB were used for chemical cell lysis in the SDS protocol and CTAB protocols, whereas commercial kits were provided with specific detergent-based lysis buffers. Based on observed DNA yields obtained by different protocols, it can be concluded that the combination of mechanical and chemical processes included in the classical protocol lysis steps were more efficient, as CPs provided higher yields of mDNA.

A drawback observed with classical protocols was that their yields were not reproducible compared with kit protocols. Reproducibility could be a result of technical variations inherent to the process itself (Salonen *et al.*, 2010). For example, the amount of collected supernatant in many of the steps (Section 2.4.1) varied between samples. It is thus reasonable to assume that variations in DNA yield are most likely to increase for extraction protocols with the most purification and wash steps (e.g. SDS and CTAB protocols used in this study). Contrastingly, with kit protocols the extracted metagenomic DNA is initially bound to a silica-based membrane and then eluted with a fixed volume. This process allowed increased uniformity in DNA yield between replicated samples.

However, and despite lower DNA yields, kits extracted better quality DNA than the classical protocols (Figure 3.4). The observed PCR-inhibition when CPmDNA was used as template could result from the presence of co-extracted plant polyphenolic and polysaccharides which are known to bind DNA and thus make it inaccessible to the polymerase enzyme (Varma *et al.*, 2007; Demeke and Jenkins, 2010; Mornkham *et al.*, 2012). It was indeed previously shown that chloroform and phenol (used in the purification steps) were not always efficient in removing plant metabolites, polyphenolic and polysaccharides (Horne *et al.*, 2004). Moreover, CTAB, SDS, phenol, chloroform and ethanol (also used in the protocols) can also contaminate extracted DNA and lead to PCR inhibition (Demeke and Jenkins, 2010).

Studies often recommend additional purification steps to eliminate plant-derived PCR-inhibitors (Demeke *et al.*, 2009; Corbisier *et al.*, 2007). These include gel electrophoresis, chromatography and use of specific chemicals (e.g. polyvinylpyrolidone, polyvinyl polypyrolidone, β-mercaptoethanol) (Sharma *et al.*, 2007; Chen *et al.*, 2009; Kit and Chandran, 2010) or enzymes (proteinases, RNase) (Demeke and Jenkins, 2010).

In the current study, PVPP columns were efficient in removing PCR-inhibitors from groundnut DNA extracted with the SDS-protocol. However, the extended purification process prolongs the extraction protocol (Sharma *et al.*, 2007), thus increasing the opportunity for technical variations (Salonen *et al.*, 2010). Commercial kits, on the other hand, were efficient in removing PCR-inhibiting compounds when compared to classical protocols as previously shown (Green *et al.*, 1999; Drabkova *et al.*, 2002).

DNA yield and quality is also dependent on plant species and/or tissue type. In this study, groundnut tissues consistently provided more DNA than sorghum or pearl millet tissues. Such observation is consistent with the one of Mace and colleagues (2003), who obtained more DNA from leaf tissues of groundnut than sorghum or pearl millet (~2:1:1 ratio, CTAB protocol). Moreover, monocotyledonous buffalo grass provided greater DNA yield than dicotyledonous cotton tissues when a classical SDS/CTAB-based protocol or Qiagen DNEasy Plant Mini Kit were used (Niu *et al.*, 2008).

The total polyphenolic and polysaccharide content differs across different plant species (Mattila and Hellström, 2007; Varma *et al.*, 2007). Groundnut tissues – including root, stems, leaves and kernels – are naturally rich in polyphenolic compounds such as ferulic acids, *p*-coumaric acid and resveratol (Chen *et al.*, 2002; Mattila and Hellström, 2007), and total phenolic content of over $100\mu m.mg^{-1}$ fresh weight has been recorded (Devi and Redi, 2002). The high concentration of phenolic compounds in the groundnut tissues therefore could explain the observed PCR inhibitions. Contrastingly, phenolic compounds from sorghum and millet root and shoot tissues have been shown to occur at lower concentrations ($50\mu m.mg^{-1}$) (Sené *et al.*, 2001).

T-RFLP was effective in showing the effects of plant DNA extracting kits on the retrievable endophytic bacterial diversity. The CTAB and the SDS protocols were shown to access the highest endophytic diversity when compared to kits with the lowest starting plant material (0.1g). When plant material was increased, the highest sorghum endophytic community diversities from roots and stems were accessed with SDS protocol with the highest reproducibility. This suggests that starting plant tissue material in DNA extractions is a limiting factor for accessibility of endophytic bacterial diversities. Contrastingly, this protocol retrieved the lowest endophytic community diversity from groundnut tissues when compared to the CTAB protocol and GeneJet kit. This is despite the additional PVPP purification process. It is apparent that co-extracted contaminants from groundnut tissues still compromised the efficiency of this protocol, thus leading to significant underestimation of groundnut-associated endophytic bacterial diversity. This further supports previous findings that plant

species rich in phenolic compounds such as groundnut require a DNA extraction protocol with a more robust purification process (Porebski *et al.*, 1997; Demeke and Jenkins, 2010). In this study, the groundnut endophytic diversity was best accessed with the CTAB-protocol.

The different protocols used in this study, with either low (0.1g) or high (0.3g) starting plant tissue material, were not able to provide complete coverage of all the OTUs observed for a specific plant tissue type. Also, protocols that retrieve high-yield DNA do not necessarily access the best diversities from plant tissue samples as shown by the groundnut community accessed with the SDS-protocol. These findings concur with previous studies that showed that there is no correlation between microbial community composition and DNA yield (Scupham *et al.*, 2007; Salonen *et al.*, 2010). Yuan *et al.*, (2012) showed that the difference in microbial structure composition of communities retrieved with different protocols is rather largely due to the lysis efficiency of mechanical and enzymatic processes involved; therefore different protocols access different bacteria. Future studies could also consider pooling mDNA samples extracted with different protocols in order to increase the range of bacterial phylotypes represented.

Other factors that could have contributed to observed community diversity differences are biases introduced by PCR and t-RFLP processes. In studies such as the current one, whereby the bacterial genomes are mixed with the host plant genome, it is impossible to measure the proportion of the bacterial genome present in the PCR reaction mixture, relative to the amount of plant DNA. Equal amounts of

mDNA in two reactions could have different proportions of bacterial and plant DNA; therefore a variation in bacterial target gene copy number (i.e. the 16S rRNA gene) is introduced (Sène *et al.*, 2001). In the current study, all controllable PCR conditions were kept constant to evenly distribute the effects of other PCR-introduced biases such as primer bias (Hansen *et al.*, 1998), accessibility of target gene sequences, target gene competition (Farrelly *et al.*, 1995) and effects of chimeric sequences (Kanawaga, 2003). These PCR-biases reduce equitable amplification of target genes found in a heterologous mixture, a phenomenon that would lead to distorted representation of indigenous communities.

However, Hartmann and Widmer (2008) showed that the highest bias in t-RFLP community studies is introduced by downstream analysis procedures. These biases could be a combination of electrophoresis efficiency and computational algorithms used to analyse peak morphology, and could lead to overestimation of peak intensity for certain T-RFs. Exclusion of short and long T-RFs as well as low intensity T-RFs for quality control purposes can also lead to exclusion of legitimate and rare bacterial phylotypes (Blackwood *et al.*, 2007). Despite these biases, t-RFLP was shown to be reliable in discerning differences introduced by experimental treatments in bacterial community communities (Hartmann and Widmer, 2008), as was also shown in this study.

The main objective in this study was to determine the most efficient DNA extraction protocol to use in order to access as much as possible the complete "endophytome" of sorghum and pearl millet plants using NGS. We conclude that the classical SDS-

based protocol constitutes is the best option for these plants with low phenolic compound content as it was able to retrieve high-yield, high PCR-quality genomic material from sorghum tissues. An increase in starting plant tissue material led to an increase in retrieved endophytic bacterial diversity.

CHAPTER 4

Identification of Endophytic Bacteria

Associated with Sorghum and Pearl Millet

Using 454 Pyrosequencing

UNIVERSITY *of the*

WESTERN CAPE

CONTENTS

UNIVERSITY *of the*

WESTERN CAPE

4.1. Introduction

Ecological studies of plant-associated endophytic bacteria provide information about the diversity, structure and functional properties of these communities. This knowledge is fundamental in understanding the establishment of endophytic communities in plant tissues, their interactions with the host plant and their responses to external environmental conditions (Hardoim *et al.*, 2008).

A diverse spectrum of endophytic bacteria – with an even broader range of functional properties – has been isolated from agricultural crops. These include PGPEBs, plant pathogens and parasites that have been shown to affect plant health and growth (Newton *et al.*, 2010). However, as previously discussed (Section 1.2), culturing techniques largely underestimate the diversity of environmental microbial communities (Rappé and Giovannoni, 2003). This consideration is more so important in the study of plant-associated endophytic bacterial communities as certain bacterial species cannot be cultured independent of the host plant or other plant-associated microbes (Aslam *et al.*, 2010). For example, a *Methylobacterium* sp. could not be re-isolated from potato plant tissues, unless it were previously co-inoculated with *Pseudomonas fluorescens* IMBG163; although the metabolic pathways responsible for this induction are not yet understood (Podolich *et al.*, 2009). Culture-independent molecular approaches such as community fingerprinting techniques and high-throughput sequencing technologies provide greater resolution in elucidating the diversity of plant-associated microorganisms (Rincon-Florez *et al.*, 2013).

The diversity of endophytic bacterial communities associated with plants that belong to the family Poaceae (includes all grasses) has been studied in great detail using both culturing and culture-independent techniques (Reinhold-Hurek and Hurek, 2011). Grasses, which constitute close to one-third of the earth's vegetative cover, are important in most ecosystems as primary food sources for herbivorous organisms (Jacobs *et al.*, 1999; Strömberg, 2011). Due to their ubiquity, and with over 10,000 known species, grasses also play a crucial role in nutrient and gaseous cycles (Kellogg, 2006; Christin *et al.*, 2009). Domesticated grasses such as maize, sorghum, rice, barley, wheat and pearl millet are cultivated to provide over 80% of food sources consumed by humans or livestock (Jacobs *et al.*, 1999; Christin *et al.*, 2009). Aesthetic and ornamental grasses including buffalo grass, bamboo, switchgrass and fountain grasses are used widely as garden art components or architectural material (King and Oudolf, 1998). Due to their ecological and economical importance, greater research efforts focus on the sustenance of the growth and health of grasses. One approach in achieving this is to examine the relationship between grasses and their associated microbial communities in order to further understand and manage plant diseases or to exploit metabolic capabilities of PGPMs for increased plant productivity (Reinhold-Hurek and Hurek, 2011).

Most studies of grass-associated endophytic bacterial communities are culture-based (Zinniel *et al.*, 2002; Magnani *et al.*, 2010; Orole and Adejumo, 2011) and/or involve characterisation of specific isolated potential PGPEBs and pathogenic bacterial strains (James and Olivares, 1997; Hurek *et al.*, 2002; Ali *et al.*, 2009; Jha and Kumar, 2009; Luna *et al.*, 2010; Luo *et al.*, 2011). Endophytic bacteria that have been isolated from grasses include plant-growth promoting hormone producers,

phosphate-solubilisers, nitrogen fixers (Luna *et al.*, 2010), parasites or pathogen antagonists (Gopalakrishnan *et al.*, 2011), soil bioremediators and toxic metal tolerant bacterial strains (Luo *et al.*, 2011).

In-depth analyses of endophytic bacterial communities using culture-independent techniques have been conducted for only a few agriculturally important grass species such as rice (Sun *et al.*, 2008; Sessitsch *et al.*, 2012), sugarcane (Magnani *et al.*, 2013) and maize (Seghers *et al.*, 2004; Pereira *et al.*, 2011). These studies have provided a solid platform for further investigation of endophytic bacteria associated with grasses. However, endophytic communities associated with other important crops in the Poaceae family such as sorghum and pearl millet have not been studied to such depth.
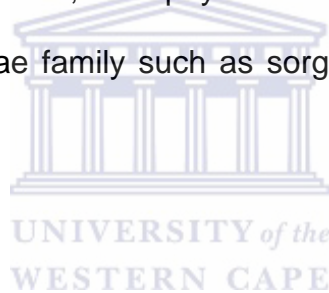
Currently, there are very few published biodiversity studies on sorghum- and millet-associated endophytic bacterial communities. In one study, culturable endophytic bacteria associated with roots of sorghum and pearl millet were affiliated to the phyla Firmicutes, Actinobacteria, Alphaproteobacteria and Betaproteobacteria (Grönemeyer *et al.*, 2012). Ramond *et al.* (2013) used t-RFLP and DGGE in a study that showed that the diversity of sorghum associated endophytic bacteria is lower than that of rhizospheric communities. However, the low resolution of the two techniques made it difficult to assign taxonomy for the recovered OTUs. To our knowledge, there is no published culture-independent study on the diversity of endophytic bacteria associated with pearl millets. Nevertheless, the effects of associations between both grasses and known PGPEBs such as *Pseudomonas* sp.

112

(AKM-P6), *Bacillus* sp. (SLS18), *Azospirillum brasilense*, *Azotobacter chroococum*; *Serratia marcescens*, *Bacillus circulans* and *Pseudomonas fluorescens* have been assessed (Tien *et al.*, 1979; Wani *et al.*, 1989; Hameeda *et al.*, 2006; Raj *et al.*, 2004; Ali *et al.*, 2009; Luo *et al.*, 2011). The infection and colonisation strategies of bacterial pathogens, for example *Herbaspirillum rubrisulbalbicans,* within sorghum and pearl millet tissues have also been studied (James *et al.*, 1997). However, the diversity of endophytic bacteria associated with sorghum and pearl millet remains largely unknown.

In the current study, high-throughput pyrosequencing was used to determine the diversity of endophytic bacterial communities associated with the roots and stems of sorghum and pearl millet. Endospheres of sorghum and pearl millet grown in South Africa are underexplored bacterial habitats. Therefore, the main aim of this study was to reveal and compare the composition of bacterial endophytes associated with these important food crops. Phylotypes with potential plant growth promoting properties are identified and considered for biofertilizer and biocontrol production as well as other industrial applications.

4.2. Results

4.2.1. Sequence data retrieved from pyrosequencing of bacterial 16S rRNA gene

A total of 135922 raw sequences were generated by pyrosequencing of the bacterial 16S rRNA genes from sorghum and pearl millet stem and root tissue metagenomic DNA. Mean raw sequence length was 383.23 ± 166.04 bp, with a minimum and maximum length of 40bp and 1177bp, respectively. After the quality trimming and filtering process using Qiime, 67016 good quality bacterial 16S rRNA gene sequences were retained for further analysis. Mean sequence length distribution of these was 400.06 ± 63.73 bp, with minimum and maximum length of 200 bp and 470 bp respectively. Each tissue sample generated an average of 16342.25 sequences (ranging from 6636 to 23442).
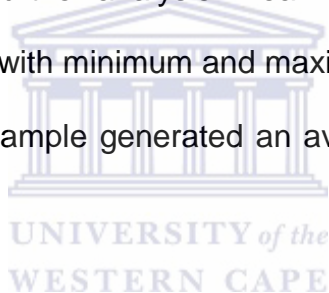
Curated sequences were initially clustered into 1146 OTUs. Preliminary taxonomic assignment of OTUs indicated that 36% of the OTUs were identified as Cyanobacteria (Streptophyta) chloroplast sequences. The sequences were manually removed and compared to sequences in the NCBI nucleotide (nt) database. These sequences had a high similarity match (≥99%) to partial chloroplast sequences of various grass species including *Sorghum bicolor* (sorghum) (Figure 4.1), *Seratia italica* (foxtail millet), *Pennisetum glaucum* (pearl millet), *Hordeum vulgares* (barley), *Zea mays* (maize) and *Triticum aestivum* (common wheat). Therefore, since they did not represent bacterial 16S rRNA gene sequences, the chloroplast sequences were manually removed from the dataset. Complete removal was not achieved; however,

OTUs affiliated to the Streptophyta chloroplast were reduced to less than 1.5% of the total OTUs.

Sorghum bicolor cultivar BTx623 chloroplast, complete genome
Sequence ID: gb|EF115542.1| Length: 140754 Number of Matches: 2

Range 1: 96063 to 96499 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
802 bits(434)	0.0	437/438(99%)	1/438(0%)	Plus/Plus

```
Query  1      GATGAACGCTGGCGGCATGCTTAACACATGCAAGTCGAACGGGAAGTGGTGTTTCCAGTG  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  96063  GATGAACGCTGGCGGCATGCTTAACACATGCAAGTCGAACGGGAAGTGGTGTTTCCAGTG  96122

Query  61     GCGAACGGGTGAGTAACGCGTAAGAACCTGCCCTTGGGAGGGGAACAACAACTGGAAACG  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  96123  GCGAACGGGTGAGTAACGCGTAAGAACCTGCCCTTGGGAGGGGAACAACAACTGGAAACG  96182

Query  121    GTTGCTAATACCCCGTAGGCTGAGGAGCAAAAGGAGAAATCCGCCCAAGGAGGGGCTCGC  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  96183  GTTGCTAATACCCCGTAGGCTGAGGAGCAAAAGGAGAAATCCGCCCAAGGAGGGGCTCGC  96242

Query  181    GTCTGATTAGCTAGTTGGTGAGGCAATAGCTTACCAAGGCGATGATCAGTAGCTGGTCCG  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  96243  GTCTGATTAGCTAGTTGGTGAGGCAATAGCTTACCAAGGCGATGATCAGTAGCTGGTCCG  96302

Query  241    AGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAG  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  96303  AGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAG  96362

Query  301    TGGGGAATTTTCCGCAATGGGCGAAAGCCTGACGGAGCAATGCCGCGTGGAGGTGGAAGG  360
              |||||||||||||||||||| ||||||||| ||||||||||||||||||||||| |||||
Sbjct  96363  TGGGGAATTTTCCGCAATGGGCGAAAGCCTGACGGAGCAATGCCGCGTGGAGGTGGAAGG  96422

Query  361    CCTACGGGTCGTCAACTTCTTTTCTCGGAGAAGAAACCAATGACGGTATCTGAGGAATAA  420
              ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Sbjct  96423  CCTACGGGTCGTCAACTTCTTTTCTCGGAGAAGAAA-CAATGACGGTATCTGAGGAATAA  96481

Query  421    GCATCGGCTAACTCTGTG  438
              ||||||||||||||||||
Sbjct  96482  GCATCGGCTAACTCTGTG  96499
```

Figure 4.1. Alignment of a representative 16S rRNA sequence identified as a Streptophyta chloroplast sequence (OTU ID 247) against the partial sorghum (*Sorghum bicolor*) BTx623 chloroplast sequence found in the NCBI nucleotide database.

The remaining sequences were re-clustered into 1036 OTUs. Distribution of the OTUs across the tissue samples is shown in Figure 4.2. The highest number of OTUs (498) was observed in the sorghum stem, whereas the pearl millet stem had the lowest number of OTUs (211). OTU sharing was observed between samples. Of

the 829 total OTUs recovered from the sorghum tissues, 17.9% were shared between the stem and root tissue samples. Pearl millet root and stem shared 17.7% of their total 664 OTUs. On average, overall OTU sharing between individual sorghum and pearl millet tissues was at 16.1(±2.6) %.
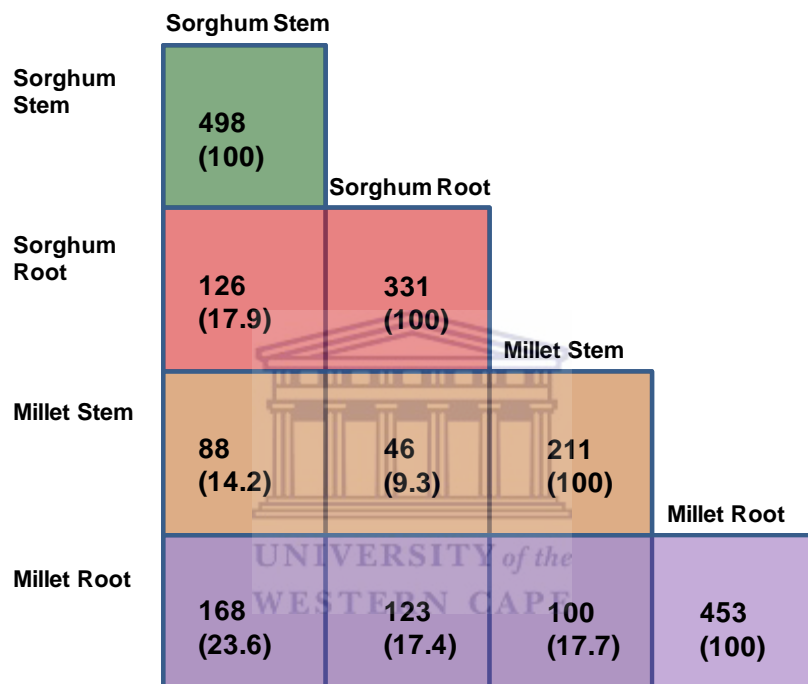
Figure 4.2. OTU sharing matrix indicates the number of shared OTUs (expressed as percentages in brackets) between tissue samples.

4.2.2. Diversity of endophytic communities found within sorghum and pearl millet tissues

Rarefaction plots were used to compare bacterial community diversities across the four plant tissue types (Figure 4.3). The pearl millet root endophytic community was the most diverse with final Chao1 value of 695.68, followed closely by sorghum stem

(645.76) and sorghum root (502.82) communities. Evenness for the three bacterial communities was approaching maximum equitability with Simpson index value of between 0.952 and 0.972. The pearl millet stem community was the least diverse (Chao1 = 360.45) and even (Simpson index = 0.857).
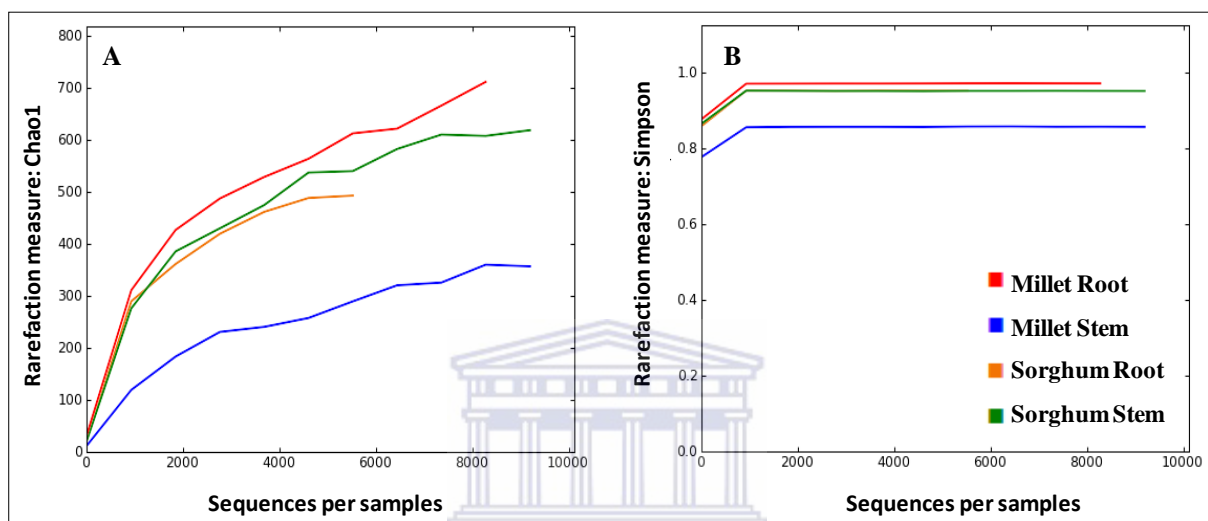
Figure 4.3. True diversity (A) and evenness (B) of endophytic bacterial communities recovered from sorghum and pearl millet tissues.

The true diversity (Chao1) rarefaction curves for bacterial communities in all tissues follow a steep increase until the sample size reaches 2000 sequences (Figure 4.3). Beyond 2000 sequences, the true diversity gradients are reduced and curves approach an asymptote as the rate at which "new" phylotypes are encountered decreases. However, a clear asymptote is not reached for all samples. This implies that "rare" species were still encountered in the last iterations; therefore it can be assumed that complete coverage of the endophytic communities was not achieved in this study (Crist and Veech 2006).

4.2.3. Structure and composition of sorghum and pearl millet endophytic bacteria communities

The phylogenetic tree created from sequences retrieved from the four tissue samples was analysed by UniFrac to determine the evolutionary relatedness of their endophytic bacterial communities. The generated PCoA plot indicated that the sorghum and pearl millet root and stem communities were distinctly different (Figure 4.4). The UniFrac F-test confirmed a statistically significant difference between each pair of communities under study (p-values ≤ 0.05).

Differences in community structure are attributable to differences in the contribution of different bacterial phyla to each community. PCoA analysis indicates that the bacterial phylotypes most responsible for observed differences are (in order of decreasing dominance) Proteobacteria, Firmicutes and Actinobacteria (Figure 4.4). Ordination of sorghum (root and stem) and pearl millet root communities was strongly influenced by the Proteobacteria and Actinobacteria phyla, whereas the pearl millet stem community ordination was most defined by the Firmicutes.

All bacterial phyla found in sorghum and pearl millet tissues are shown in Figure 4.5. This figure confirms findings in Section 4.2.2. that endophytic communites found in both sorghum tissues and pearl millet root are diverse. These tissues were dominated by Proteobacteria (78.27% - 85.34%), Firmicutes (3.33% - 7.16%), Actinobacteria (4.36% - 12.64%), Cyanobacteria (0.31% - 1.66%) and Bacteroidetes (0.79% - 1.95%) (Figure 4.5). Pearl millet stem community was the least diverse as it

was largely dominated by Firmicutes (82.89%), Proteobacteria (13.02%) and Cyanobacteria (2.45%). Less than 2% of the pearl millet stem community was made up of Actinobacteria and Bacteroidetes. It is clear that bacterial phyla that were shown to contribute the most to differences between the four communities (Figure 4.4), are in fact, the most dominant groups in the respective tissues.
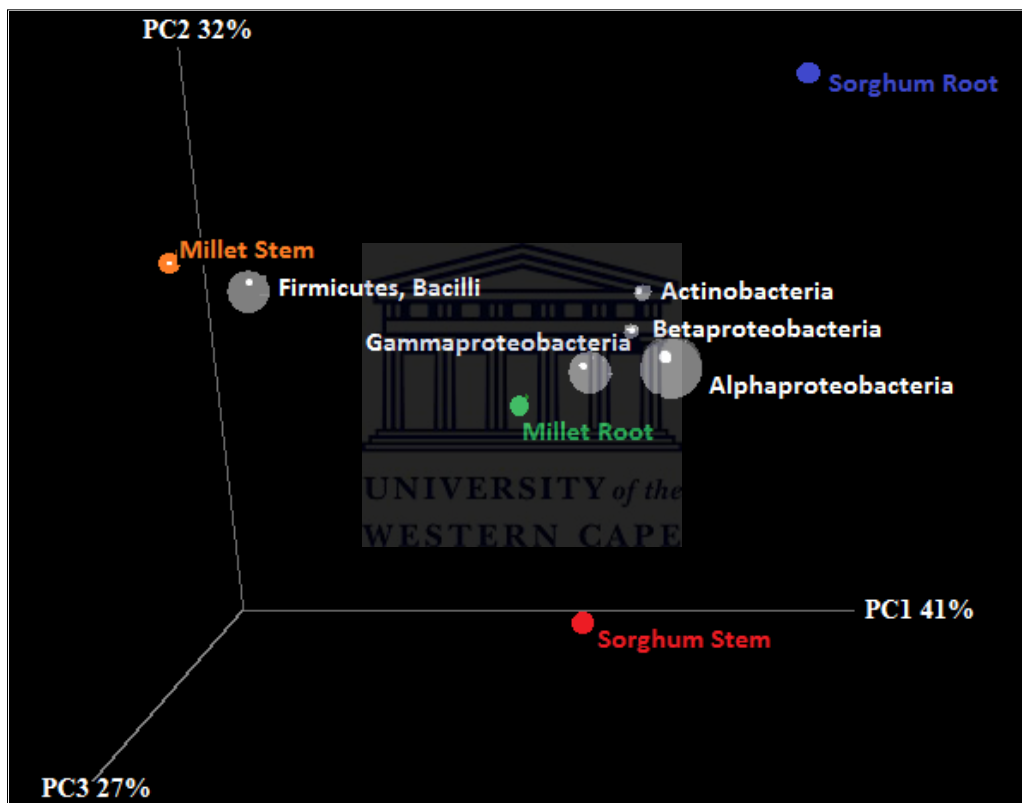
Figure 4.4. Three-dimensional UniFrac PCoA ordination of endophytic bacteria communities recovered from sorghum and pearl millet tissues. Grey bubbles represent major bacterial taxa that contribute to observed variations (size of bubble is proportional to dominance and position to principal component most influenced).
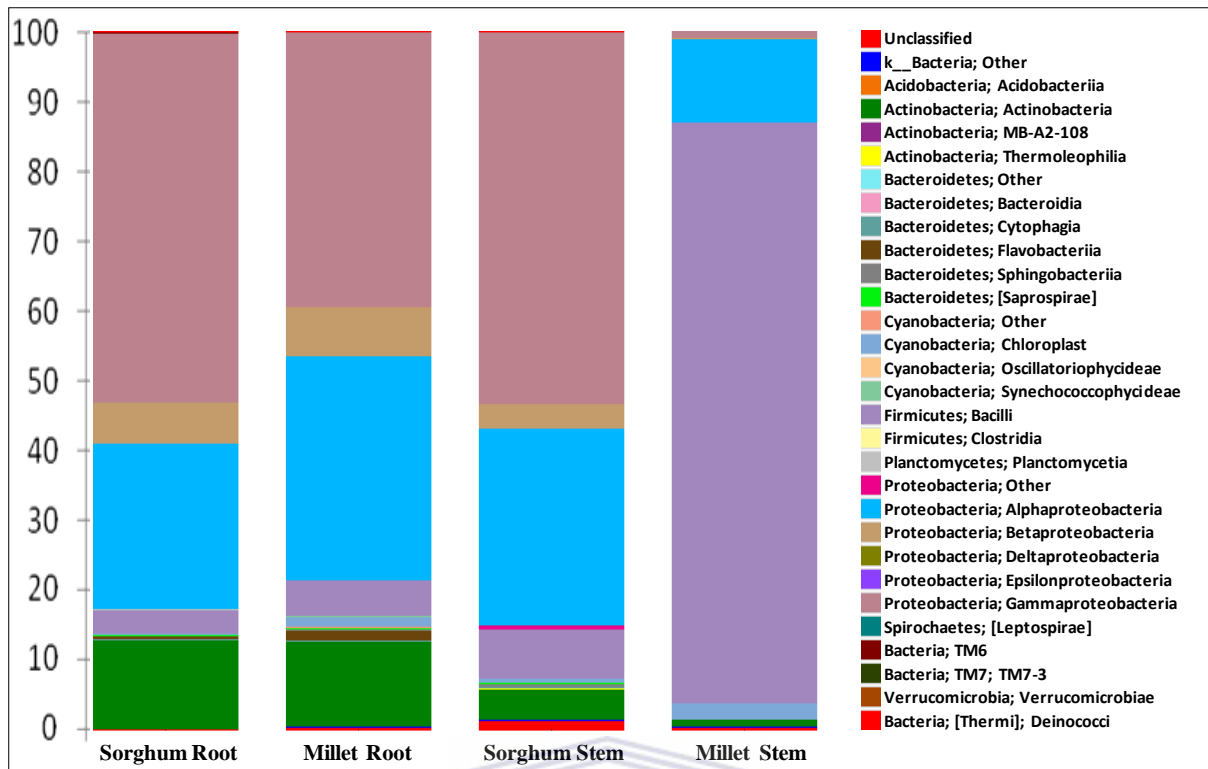
Figure 4.5. Relative abundance of major bacterial lineages recovered from sorghum and pearl millet stem and root tissues. Relative abundance is calculated as the percentage of sequences belonging to a particular lineage out of all 16S rRNA gene sequences recovered from a given plant tissue type.

A total of 112 genera were identified in this study, and they were affiliated to the phyla Proteobacteria (59), Actinobacteria (27), Firmicutes (10), Bacteroidetes (9) and Cyanobacteria (2). Other genera – these include representatives of Planctomycetes, Thermi, Verrucomicrobia, TM-6 and TM-7 – constituted less than 1% of all observed phylotypes.

It is important to note that often, during taxonomic affiliation of OTUs, more than one OTU was assigned to a specific bacterial phylotype. For example, more than ten OTUs were assigned to genera such as *Microbacterium*, *Leuconostoc*, *Erwinia*, *Pseudomonas* and *Stenotrophomonas*. Some of these OTUs (derived from one or more tissues) could belong to the same bacterial species and/or strain; however it is also possible that the different OTUs represent different bacterial species and strains under the same genus. Therefore, Figure 4.6 shows that a greater proportion (59.8%) of bacterial phylotypes occurred in two or more tissues, than previously suggested (Figure 4.2). The sorghum stem and root tissues shared 39.6% of their total combined identified phylotypes, whilst 28% of the total pearl millet endophytic community was shared between the stem and root tissues. Bacterial phylotypes shared between sorghum and millet tissues constituted 40.2% of the total phylotypes observed in this study.
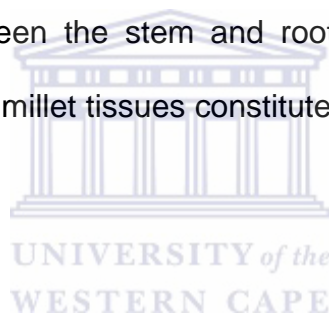
Differences in relative abundance of specific genera between all pairs of communities were also observed. For example, Proteobacteria had a relatively high representation in all four tissues, accounting for 64% of total observed phylotypes. Gamma-(γ)-proteobacteria was the most dominant class in both sorghum tissues and pearl millet root, contributing 39.2% and 53.1%, respectively. The dominating γ-proteobacteria genus in sorghum root and pearl millet root was *Pseudomonas* (46.54% and 16.3%, respectively), whereas *Erwinia*, an enterobacteria, dominated the sorghum stem community at 30.2%. Contrastingly, γ-proteobacteria only made up 0.9% of the pearl millet stem community, where Alpha(α)-proteobacteria (11.9%) were more dominant.

Community diversity estimations were based on the analysis of single metagenomic DNA samples from each individual tissue. In the absence of replicated data for comparative analysis, it was impossible to assess the bias introduced by experimental error. Also, there was no parallel analysis of the communities found in the most likely contaminating environment, that is, the rhizospheric soils or phylosphere environments of these plants. Therefore, this study cannot determine, conclusively, the bacterial phylotypes that are enriched in the individual tissues under study. Nevertheless, it was previously shown that recovery of dominant phylotypes is more reproducible than that of rare phylotypes (Charlson *et al.*, 2012). Therefore, to exercise caution, discussions on potential metabolic properties of communities found in this study are restricted to the 25 most dominant phylotypes (≥1% of the total population in either one of the tissues) (Table 4.2).
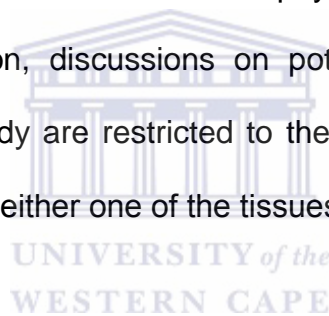
Dominant genera were from the phyla Actinobacteria, Bacteroidetes, Firmicutes and Proteobacteria (Table 4.1). The most dominant Actinobacteria were from the *Microbacterium* genus, and these were particularly dominant in the roots of sorghum (9.7%) and pearl millet (3.4%). *Curtobacteria* were dominant in sorghum roots (1.6%) and stem (2.0%), whilst *Rhodococci* were more restricted to sorghum stem (1.5%). *Arthrobacter* were only observed in pearl millet roots (3.2%). The genus *Chryseobacterium* was the only dominant representative of the Bacteroidetes, and it constituted just up to 1% of the root communities of both plants.

Five representative genera from the Firmicutes phylum were considered dominant. Two of these genera, *Paenibacillus* and *Bacillus* occurred in all tissues, albeit at comparatively lower abundance. *Exiguobacteria* were restricted to pearl millet roots (3.9%), *Leuconostocs* to pearl millet stem (82.8%) and *Lactococci* to sorghum stem (6.7%). Notable dominant Proteobacteria included *Agrobacteria* (7.3%), *Rickettsiales* (9.4%), *Erwinia* (9.2%), *Pseudomonads* (18.4%) and *Strenotrophomonads* (6.0%). All five of these occurred in both sorghum and pearl millet. Other Proteobacteria were found in lower abundance. *Rhizobia* (1%) and *Sphingobium* (4.3%) were only dominant in sorghum roots. *Swaminathania* (2.4%) and *Sphingomonas* (2.7%) were restricted to the sorghum stem tissue, whilst *Methyloversatilis* (1.9%) and *Janthinobacteria* (2.0%) were predominant in pearl millet root. *Herbaspirillum* were predominant in sorghum roots (3.4%) and stem (3.6%).
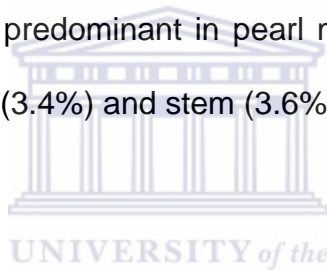
Some OTUs could only be identified to the order or family level, and the genus was unclassified, e.g., Rhizobiales, Rickettsiales and Xanthomonadaceae. These OTUs mostly matched partial sequences from environmental samples. Unclassified and unidentified bacteria constituted 0.9% of total observed OTUs. These OTU sequences were manually isolated and matched to sequences in the NCBI nucleotide database. The majority of sequences matched partial mitochondrial and chloroplast sequences of grass species and partial 16S rRNA sequences of previously uncultured bacterial clones (Table 4.2). Some sequences closely matched 16S rRNA sequences of known bacterial species. Sequence similarity match was less than 97% for all alignments.

Acidobacteria: Terriglobus **Actinobacteria:** Cryocola, Nakamurellaceae sp., Nocardaceae sp., Pimelobacter, Williamsia, Patulibacteraceae sp., Patulibacter, Sphingobacterium, **Cyanobacteria:** unidentified Cyanobacterium. **Firmicutes:** Bacilllaceae, Paenibacillaceae sp., Enterococcaceae sp., Streptococcaceae sp., lactococcus, Clostridiceae sp., Clostridium. **Proteobacteria:** Pleomorphomonas, Phylobacteriaceae sp., Ancylobacter, Gluconobacter, Pigmentiphaga, Curvibacter, Pelomonas, Yersinia, Coxiellaceae sp., Luteibacter, **Spirochaetes:** Leptospira

Actinobacteria: Mycetocola. **Bacteroidetes:** Bacterodes, Spirosoma, Chitinophagaceae sp., Sediminibacterium. **Proteobacteria:** Phenylobacter, Roseomonas, unspecified Burkholderiales sp., Acidovorax **Thermi:** Thermus

Actinobacteria: Brachybacterium, Geodermartophilaceae sp., Intrasporangiaceae sp. Knoellia, Phycicoccus, Agromyces, Leucobacter, Rathayibacter, Micrococcaceae sp., Arthrobacter, Micromonosporaceae sp., Actinoplanes, Sptreptomycetaceae sp., MB-A2-108, Solirubrobactericeae sp., **Bacteroidetes**: Bacteroidetes sp., Cytophagagceae sp. **Cyanobacteria:** Chroococcales sp., Phormidium, Acaryochoccus. **Firmicutes:** Planococcaceae sp., Staphylococcus, **Proteobacteria:** BD7-3 sp., Bradyrhizobiceae sp., Balneimonas, Acetobacteraceae sp., Roseococcus, Phaeospirillum, Betaproteobacterium sp., Achromobacter, Neisseriaceae sp., Hydrogenophilus, Alteromonas, Shewanella, Enterobacter sp., Lysobacter, Xanthomonas, **TM7-3**

Actinobacteria: Mycobacterium. **Proteobacteria:** Aurantimonadeceae sp., Kaista, Xanthobacteraceae sp., Acetobacteraceae sp., Sphingomonadales sp., Burkholderia, Comamonas, Gammaproteobacterium sp., Trabulsie

Actinobacteria: Actinomyces, Pseudonocardia, **Cyanobacterium:** unidentified cyanobacterium. **Firmicutes:** Bacilli sp, Leuconostoccaceae sp., Proteobacteria: Caulobacter, Pseudoruegeria, Sulfurospirillum, Reinekea

Actinobacteria: Microbacteriaceae sp., Salinibacterium, Kocuria, Micrococcus, Norcadiodes. **Bacteroidetes:** Dyadobacter, Sphingobacteriaceae sp., Chitinophagaceae sp., Chitinophaga, **Firmicutes:** Bacillales sp., Clostridia sp. **Proteobacteria:** Bradyrhizobium, Devosia, Rhodoplanes, Rhizobium, Rhodospirillaceae sp., Alcaligenaceae sp., Leptothrix, Ramlibacter, Haliangiaceae sp., Legionellaceae sp., Legionella, **TM7-3**, **Verrucomicrobia**: Luteolibacter. **Thermi**: Deinococcus

Bacteroidetes: Polaribacter. **Proteobateria:** Loktanella, Alternomonadaceae sp., Psychromonas, Xanthomonadaceae sp.

Actinobacteria: Actinomycetales sp., Corynebacterium, Rhodococcus. **Proteobacteria:** Caulobacterceae sp., Rhodobacteraceae sp., Magnetospirillum, Xanthomonadaceae sp.

Actinobacteria:Aeromicrobium. **Proteobacteria:** Swaminathania, Blastomonas

Bacteroidetes: Chryseobacterium, Pedobacter. **Firmicutes:** Bacillaceae sp., Exiguobacterium, Alkalibacterium. **Proteobacteria:** Bosea, Rhizobiaceae sp., Paracoccua, Polaromonas, Variovorax., Oxalobacteraceae sp., Herbaspirillum, Janthinobacterium., Enterobacteriacea sp., Pseudomonadaceae sp., Stenotrophomonas. **TM6 sp.**

Unclassified bacteria. Actinobacteria: Microbacteriaceae sp., Curtobacterium, Microbacterium. **Cyanobacteria:** Unspecified **Firmicutes:** Bacillus, Paenibacillus. **Proteobacteria:**Unclassified, Caulobacteriaceae sp., Methylobacteriaceae sp., Methylobacterium, Agrobaccterium, Rickettsiales, Sphingomonadaceae sp., Sphingomonas, Comamonadaceae sp., Cupriavidus, Ralstonia, Methyloversatilis, Erwinia, Pseudomonas, Vibrionaceae sp.

Firmicutes: Brevibacillus, Leuconostoc. **Proteobacteria:** Novosphingonium

Actinobacteria: Brevibacterium., Micrococcaceae sp., Nocardioidaceae sp. **Bacteroidetes:** Flavobacterium, Weeksellaceae sp. **Firmicutes:** Geobacillus. **Planctomycetes:** Isophaeraceae sp. **Proteobacteria:** Bradyrhizobiaceae sp., Sphingomonaceae sp., Sphingobium, Idiomarina

Proteobacteria: Unspecified α-proteobacteria

Actinobacteria: Propionibacterium

Sorghum Stem 28
Millet Root 38
Sorghum Root
Millet Stem
26 10 11 5 9
17 7
26
11 3
1 3
1

Figure 4.6. Venn diagram representation of the composition of sorghum and pearl millet endophytic bacterial communities and distribution of specific phylotypes across the different plant tissues.

Table 4.1. Relative abundance (bold) of the twenty-five most dominant bacterial phylotypes found in sorghum and pearl millet tissues. Total number of OTUs assigned to each specific phylotype is shown in brackets.

Phylum/Class	Order	Family	Genus	Total	Sorghum Root	Sorghum Stem	Millet Root	Millet Stem
Actinobacteria	Actinomycetales	Microbacteriaceae	Curtobacterium	**1.3** (7)	**1.6** (4)	**2.0** (7)	**0.8** (7)	**0.6** (5)
Actinobacteria	Actinomycetales	Microbacteriaceae	Microbacterium	**3.3** (10)	**9.7** (8)	**0.2** (3)	**3.4** (6)	**0.1** (3)
Actinobacteria	Actinomycetales	Micrococcaceae	Arthrobacter	**0.8** (2)	**0.0**	**0.0**	**3.2** (2)	**0.0**
Actinobacteria	Actinomycetales	Nocardiaceae	Rhodococcus	**0.4** (4)	**0.0**	**1.5** (4)	**0.2** (2)	**0.0** (1)
Bacteroidetes	Flavobacteriales	Weeksellaceae	Chryseobacterium	**0.4** (5)	**0.3** (2)	**0.0** (1)	**1.1** (3)	**0.0**
Firmicutes (Bacilli)	Bacillales	Bacillaceae	Bacillus	**0.4** (4)	**1.2** (4)	**0.0** (2)	**0.3** (3)	**0.0** (1)
Firmicutes (Bacilli)	Bacillales	Paenibacillaceae	Paenibacillus	**0.7** (6)	**1.8** (2)	**0.1** (2)	**0.8** (2)	**0.1** (1)
Firmicutes (Bacilli)	Bacillales	Exiguobacteraceae	Exiguobacterium	**1.0** (4)	**0.0**	**0.0**	**3.9** (4)	**0.0**
Firmicutes (Bacilli)	Lactobacillales	Leuconostocaceae	Leuconostoc	**20.7** (19)	**0.1** (1)	**0.0**	**0.0**	**82.8** (19)
Firmicutes (Bacilli)	Lactobacillales	Streptococcaceae	Lactococcus	**1.7** (7)	**0.0**	**6.7** (7)	**0.0**	**0.0**
Alphaproteobacteria	Rhizobiales	Unclassified	Unclassified	**0.6** (10)	**1.6** (6)	**0.5** (5)	**0.4** (6)	**0.0** (1)
Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Unclassified	**1.3** (2)	**0.2** (1)	**0.2** (1)	**4.9** (1)	**0.0**
Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Agrobacterium	**7.3** (6)	**8.7** (4)	**14.2** (4)	**6.3** (4)	**0.0**
Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Rhizobium	**0.3** (1)	**1.0** (1)	**0.0**	**0.0**	**0.0**
Alphaproteobacteria	Rhodospirillales	Acetobacteraceae	Swaminathania	**0.6** (9)	**0.0**	**2.4** (9)	**0.0**	**0.1** (5)
Alphaproteobacteria	Rickettsiales	Unclassified	Unclassified	**9.4** (18)	**3.5** (8)	**4.5** (9)	**18.8** (15)	**11.0** (15)
Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingobium	**1.1** (1)	**4.3** (1)	**0.0**	**0.0** (1)	**0.0**
Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas	**0.8** (8)	**0.3** (2)	**2.7** (7)	**0.1** (2)	**0.0**
Betaproteobacteria	Burkholderiales	Oxalobacteraceae	Herbaspirillum	**1.6** (7)	**3.4** (4)	**2.6** (7)	**0.2** (4)	**0.0**
Betaproteobacteria	Burkholderiales	Oxalobacteraceae	Janthinobacterium	**0.6** (5)	**0.3** (2)	**0.1** (1)	**1.9** (4)	**0.0**
Betaproteobacteria	Rhodocyclales	Rhodocyclaceae	Methyloversatilis	**0.6** (4)	**0.1** (3)	**0.4** (4)	**2.0** (4)	**0.1** (3)
Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Erwinia	**9.2** (15)	**0.1** (5)	**30.2** (15)	**6.6** (15)	**0.1** (3)
Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas	**18.4** (34)	**46.5** (32)	**10.1** (29)	**16.3** (30)	**0.7** (9)
Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Unclassified	**2.0** (7)	**0.0**	**0.0**	**7.9** (7)	**0.0** (1)
Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Stenotrophomonas	**6.0** (25)	**4.7** (10)	**11.5** (23)	**7.9** (12)	**0.0**

Table 4.2. Match results of unidentified OTU sequences to sequences in the NCBI nucleotide database using Blastn tools.

Number of sequences	Closest matching NCBI sequence	Accession number	Sequence source
683	*Sorghum bicolor* mitochondrion	DQ984518.1	Sorghum root and stem
894	*Tripsacum dactyloides* (gamagrass) mitochondrion	DQ984517.1	All tissues
1	*Arachis hypogaea* 19S ribosomal RNA gene	EU307403.1	Sorghum stem
1	*Puelia olyriformis* (grass species) mitochondrion	HQ604062.1	Sorghum stem
4	*Eleusine coracana* (finger millet) mitochondrion	HQ183502.1	Millet root and stem
1	*Flagellaria indica* (whip vine) mitochondrion	HQ183503.1	Millet stem
1	Alpha proteobacterium 10819, 16S rRNA gene	EF422209.1	Sorghum root
5	*Pseudomonas mendocina* ymp, 16S rRNA gene	CP000680.1	Millet and sorghum stems
12	*Pseudomonas putida* GB-1, 16S rRNA gene	CP000926.1	Millet stem and root
1	*Leuconostoc palmae* TMW 2.694 16S rRNA gene	AM940225.1	Millet stem
1	*Pseudomonas fluorescens* Pf0-1, 16S rRNA gene	CP000094.2	Sorghum root
9	*Pantoea vagans* C9-1 chromosome	CP002206.1	Sorghum stem
199	*Pseudomonas putida* BIRD-1, 16S rRNA gene	CP002290.1	All tissues
1	Proteobacteria 16S rRNA, from uranium mine soil	HQ706442.1	Sorghum stem
6	*Pseudomonas brassicaceurum* 16S rRNA gene	CP002585.1	sorghum and millet root
1	*Magnetospirillum* sp. pMbN1, 16S rRNA gene	FQ790395.1	Sorghum stem
1	Alpha proteobacterium CC-2 16S rRNA gene	JF490043.1	Sorghum stem
12	*Lactococcus lactis* strain PSY2, 16S rRNA gene	JF703669.1	Sorghum stem
221	Uncultured bacterial clones, 16S rRNA gene	Various	All tissues

4.2.4. Metabolic potential of sorghum and pearl millet endophytic communities

Ten metabolic properties of endophytic bacteria were selected for the characterisation of dominant genera associated with sorghum and millet. These include agriculturally relevant traits such as plant growth promoting and phytopathogenic properties, as well as metabolic properties (non-exhaustive) that are often targeted for industrial applications.

It is important to note that one genus often comprises of more than one species, and the collective properties of each species are unique. Therefore, in this characterisation, published peer-reviewed articles were surveyed to determine if each individual genus has within it at least one bacterial species that exhibits either of the characteristics shown in Figure 4.7.

GENUS	SR	SS	MR	MS
Curtobacterium	•	•	•	•
Microbacterium	•	•	•	•
Arthrobacter			•	
Rhodococcus		•	•	
Chryseobacterium	•	•	•	
Bacillus	•	•	•	
Paenibacillus	•	•	•	•
Exiguobacterium			•	
Leuconostoc	•			•
Lactococcus		•	•	
Agrobacterium	•	•	•	
Rhizobium	•	•		
Swaminathania	•			•
Sphingobium	•		•	
Sphingomonas	•	•	•	
Herbaspirillum	•	•	•	
Janthinobacterium	•	•	•	
Methyloversatilis	•	•	•	•
Erwinia	•	•	•	•
Pseudomonas	•	•	•	•
Stenotrophomonas	•	•	•	

Key (METABOLIC ACTIVITY): Phytohormone prod.; N$_2$ fixation; P solubilisation; Siderophore prod.; Plant polymer degrad.; Lipase production; Xenobiotic degrad.; Metal tolerance; Antimicrobial; Human pathogen; Plant pathogen

Figure 4.7. Functional characterisation of sorghum and pearl millet root and stem endophytic communities. Colour codes are used to highlight genera that have representative species exhibiting specific characteristics, and white space shows that no representative species were found on record. Black dots indicate presence/absence of the genus in the corresponding plant tissue.

Dominant phylotypes within sorghum and pearl millet root and stem tissues have the potential for a broad range of metabolic activities (Figure 4.7). All genera with the exception of *Leuconostoc*, *Janthinobacterium* and *Methyloversatilis* were all shown to have representative species with plant growth promoting properties, including those involved in increasing the host plant's biomass (production of plant growth promoting phytohormones) and its access to nutrients (nitrogen fixation, siderophore production, phosphate solubilisation). Most of these potential plant growth promoting genera, i.e. *Curtobacterium*, *Microbacterium*, *Bacillus*, *Paenibacillus*, *Erwinia* and *Pseudomonas*, were detected in all four tissues. Other metabolic properties of biotechnological interest (lipase production, metal tolerance, xenobiotic degradation and plant polymer degradation) were well represented in most of the observed genera. Figure 4.7 shows that there is little information in literature regarding the metabolic potential of genera *Swaminathania*, *Janthinobacteria* and *Methyloversatilis*. It is possibly that genera with no type species exhibiting a specific property could have within it such a species, although still uncharacterised in that regard, or yet to be discovered. Therefore, this analysis is only exploratory as conclusive evidence can only be based on culture-based analysis or survey of specific genes responsible for each individual property.

4.3. Discussion

The current study is the first to report on the diversity of endophytic bacterial communities associated with sorghum and pearl millet as revealed by pyrosequencing analysis. Deep analysis of the sorghum and pearl millet endophytome shows that these South African crops are associated with a diverse spectrum of bacterial phylotypes including representatives of Firmicutes, Proteobacteria, Actinobacteria, Cyanobacteria and Bacteroidetes. These bacterial phyla were previously noted to be prevalent in plant tissue environments (Rosenblueth and Martínez-Romero, 2006; Reinhold-Hurek and Hurek, 2011).
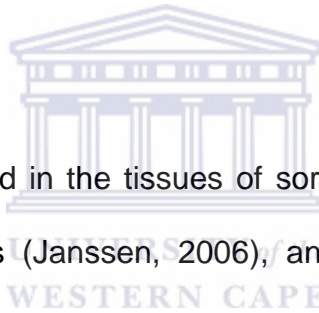
Dominant bacterial phyla found in the tissues of sorghum and pearl millet are also common in soil environments (Janssen, 2006), and were frequently detected in previous culture-based studies of soil bacterial communities (Figure 4.8). Therefore, it is possible that pearl millet and sorghum endophytic bacteria were mostly recruited from the surrounding soil environment. On the other hand, other well-known soil bacterial phyla such as Acidobacteria, Verrucomicrobia and Planctomycetes occurred in very low abundance in sorghum and pearl millet tissues. Chloroflexi and Gemmatimonadetes bacteria were not detected at all in these tissues. This could suggest that these groups are present in low abundance in the surrounding soil. For instance, the abundance of Verrucomicrobia – only detected in sorghum roots in this study (Figure 4.6) – in the soil has been shown to decrease significantly with cultivation and decreasing soil moisture content (Buckley and Schmidt, 2001). Therefore, it is possible that the abundance of Verrucomicrobia in the soil and plant

tissues were affected by the below average rainfall prior to harvest (SAWS, 2012). Alternatively, the low abundance of certain soil microorganisms in sorghum and pearl millet tissues could be due to the selectivity of the recruitment process in the roots, which would be biased towards specific phylotypes (Hardoim *et al.*, 2008).
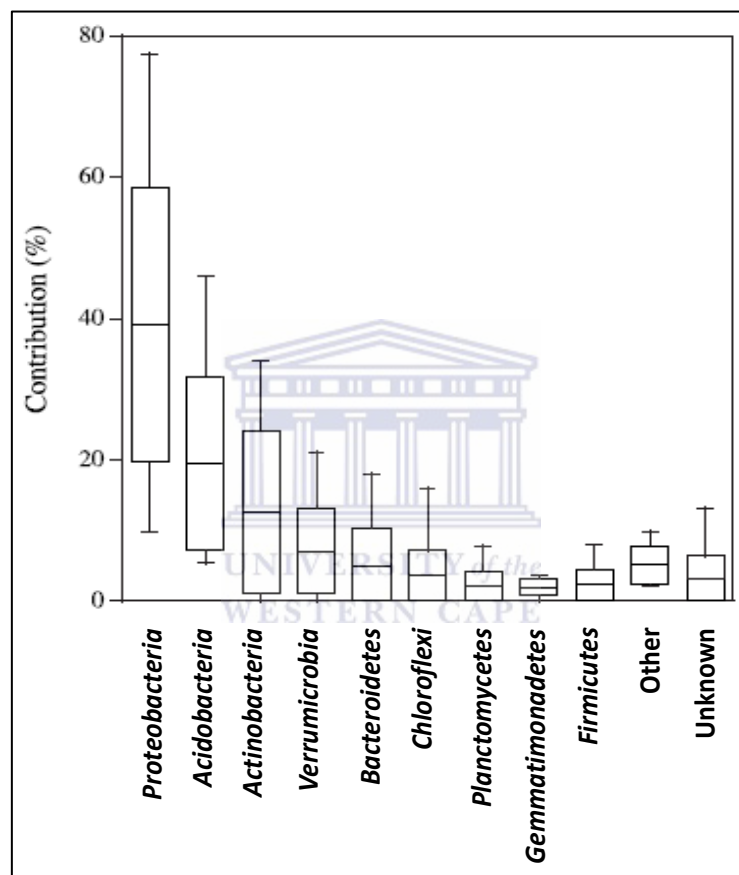
Figure 4.8. Contributions of 16S rRNA and 16S rRNA genes from members of different phyla in libraries prepared from soil bacterial communities (2920 clones in 21 libraries). The horizontal line in the middle of each block indicates the mean, the block represents 1 standard deviation on either side of the mean, and the vertical lines extending above and below each block indicate the minimum and maximum contributions of each phylum (Adapted from Janssen, 2006).

Soil bacterial communities from which endophytic bacteria are recruited are, in turn, affected by the soil conditions and farming practices (Buckley and Schmidt, 2001). Plants used in this study, sorghum and pearl millet, were cultivated using different techniques. The sorghum field was treated with two rounds of fertilization and one pesticide spray with supplementary irrigation. Rotation farming was also practised on this field, whereby sorghum crop was alternated with sunflower. Pearl millets, on the other hand, were grown tillage- and irrigation-free with one round of fertilization in a field where rotation farming was not practised. It can be argued that the different farming techniques affected the soil structure, chemistry, water retention potential and nutrient content; factors which were previously shown to affect the structure of rhizospheric communities (Fierer *et al.*, 2003; Fierer and Jackson 2005; Ramond *et al.*, 2013; Hansel *et al.*, 2008). This could have contributed to the observed differences in the structure of sorghum and pearl millet endophytic communities.
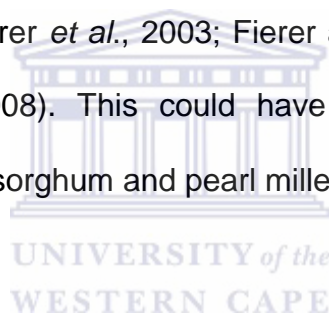
Pearl millet stem tissue had a high abundance of Leuconostocs (>80%). Leuconostocs are hetero-fermentative bacteria known to infect agricultural crops during harvest (Watt and Cramer, 2009). This implies that the pearl millet stem tissues were not healthy at harvest, despite the lack of visible symptoms. Infection of the pearl millet stem tissue could have occurred through surface lesions or wounds prior to harvest. Aphids, which were observed on the pearl millet field, could also be responsible for this infection as they are known to be effective transmitters of phytopathogens (Goggin, 2007). Alternatively, infection could have taken place during or after sampling, thus indicating that sample handling and/or transportation processes were not adequately aseptic.

Analysis of the endophytic communities in this study was carried out using one metagenomic DNA sample from each individual tissue. The lack of replicate samples or parallel analysis of soil communities makes it impossible to determine with certainty the core bacterial communities associated with sorghum and pearl millet tissues, or the proportion of bacteria that were specifically recruited from the soil and enriched in these plant tissues (Charlson *et al.*, 2012). Therefore, these aspects should be considered in future studies of sorghum and pearl millet endophytic communities using NGS. Other approaches could include analysis of temporal/seasonal and spatial (geographic) shifts of these bacterial communities, as well as the effects of different environmental factors on their structure (Fierer *et al.*, 2003; Fierer and Jackson, 2005; Ramond *et al.*, 2013). These studies would contribute towards revealing specific bacterial groups that are permanently associated with specific plants, and those that are integral to the plant's response to environmental stress.
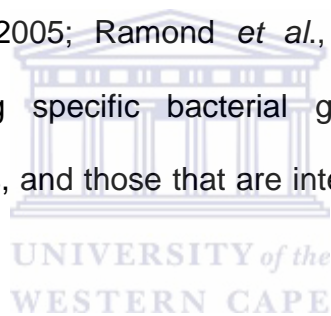
Nevertheless, and despite the lack of replicate analysis, this study was sufficient to show that sorghum and pearl millet tissues harbour diverse bacterial taxa of biotechnological significance. Most dominant bacterial genera observed in sorghum and pearl millet (Table 4.1) represent some previously characterised bacteria that were isolated from other graminaceous and non-graminaceous plants such as maize, sugarcane, rice, poplar, grapevine and sunflower (Pereira *et al.*, 2011; Magnani *et al.*, 2010; Sun *et al.*, 2008; Ulrich *et al.*, 2008, Compant *et al.*, 2011; Ambrosini *et al.*, 2012).

Most genera observed in sorghum and pearl millet plants, e.g., *Curtobacterium*, *Microbacterium*, *Bacillus*, *Paenibacillus*, *Agrobacterium*, *Chryseobacterium*, *Sphingomonas*, *Herbaspirillum*, *Erwinia*, *Pseudomonas* and *Stenotrophomonas*; have been implicated in at least one direct plant growth promoting activity (Tsavkelova *et al.*, 2006; Franche *et al.*, 2009). For example, *Chryseobacterium*, *Sphingobacterium* and *Ralstonia* species isolated from maize rhizosphere were shown to produce plant growth promoting IAA and siderophores, and subsequent *in vitro* inoculation of the maize rhizosphere with these bacteria led to significantly increased plant biomass (Marques *et al.*, 2010). Whilst *Chryseobacteria* were dominant in this study (0.4%), most notably in pearl millet root (1.1%), *Ralstonia* occurred at very low levels (0.03-0.2%) in all tissues (Figure 4.6). *Sphingobacteria* were only found in the sorghum stem (Figure 4.6), also at very low abundance (0.5%). Other bacterial genera also previously shown to produce plant growth promoting hormones include *Bacilli*, *Pseudomonads* and *Rhizobia* (Matiru and Dakota 2004; Luo *et al.*, 2011; Reis *et al.*, 2011).

Endophytic bacteria with potential to increase the plant's access to nutrients were dominant in sorghum and pearl millet tissues. Diazotrophic bacteria are able to fix atmospheric nitrogen into forms that are usable by the bacteria and the host plant. The order Rhizobiales in particular, includes nitrogen-fixing bacteria that are known for their symbiotic relationships with leguminous plants (Kraizer *et al.*, 2011). However, grass species have also been shown to benefit from their association with nitrogen-fixing bacteria, either through direct assimilation of fixed nitrogen or from growth stimulating phytohormones produced by these bacteria (Reis *et al.*, 2011). In fact, free-living diazotrophic bacterial strains such as the *Burkholderia* and

Herbaspirillum species are specifically targeted for biofertilization of non-leguminous crops including rice, maize, sugarcane and sorghum (James and Olivares, 1997; Baldani *et al.*, 2000; Salles *et al.*, 2004; Bhattacharjee *et al.*, 2008). Another interesting genus is *Swaminathania*, which has metabolically versatile species such as the diazotrophic, salt-tolerant *Swaminanathania salitolerans* with P-solubilising capabilities (Loganathan and Nair, 2004).

The high abundance of *Pseudomonads* in sorghum and pearl millet tissues (Table 4.1) strongly suggests the presence of bacteria-mediated phosphate-solubilisation processes within these plants (Kuklinsky-Sobral *et al.*, 2004). Several *Pseudomonas* species (e.g., *P. putida*, *P. capsis*, *P. flourescens*, *P. aeruginosa*) have been shown to use organic acids such as 2-ketogluconate, gluconate, tartaric acid, formic acid and acetic acid to solubilise inorganic phosphates and make phosphorus available to plants (Park *et al.*, 2009; Miller *et al.*, 2010). Other well-known phosphate solubilising bacteria that were dominant in sorghum and pearl millet tissues include species from the genera *Bacillus*, *Rhizobium*, *Agrobacterium* and *Erwinia* (Rodriguez and Fraga, 1999).

In this study, *Pseudomonads* occurrence appeared to be greater in the roots (46.5% and 16.3% in sorghum and pearl millet roots, respectively) than in the stem tissues. *Pseudomonads* are prevalent in rhizospheric soils (Lugtenberg *et al.*, 2001). It is likely that the high abundance in the roots was due to increased horizontal transmission of these bacteria from the soil into the roots. Like most competent endophytes, *Pseudomonas* species possess traits that enable them to compete

successfully in the endosphere. These include chemotactic features such as surface receptors and motility structures (flagella and pili) that allow the rhizospheric bacteria to recognise specific plant exudate compounds and migrate towards the host plant (De Weert *et al.*, 2002). *Pseudomonads* can also secrete a mucilagenous substance that allows cells to attach to the root and form microcolonies (Chin-A-Woeng *et al.*, 1997). Endophytic *Pseudomonads* enter plant roots through lesions on the root surfaces as well as through the root hairs (Prieto *et al.*, 2011), and are then able to quickly migrate to other plant tissues using their motility features. Figure 4.9 depicts the efficiency with which *Pseudomonas* strains are able to colonize plant roots.
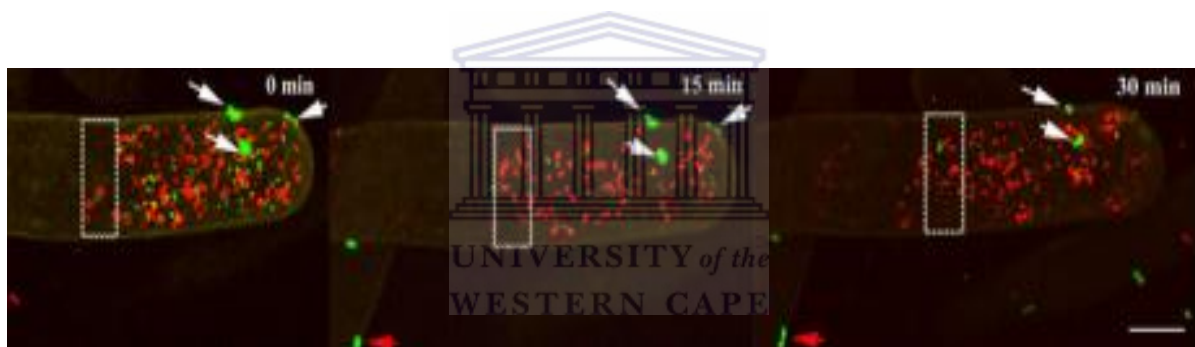
Figure 4.9. Simultaneous colonization of in vitro olive cv. Manzanilla root hairs by *Pseudomonas* PGPB strains, *P. putida* PICP2 (green, EGFP-tagged) and *P. flourescens* PICF7 (red, RFP-tagged) at three time-points (rectangular inset used as reference to assess bacterial movement). At t=0, high influx of PICF7 (predominant) and PICP2 cells into the root tissue is observed, with several cells attached to the root surface (white arrow). Displacement of cells from the root hair tip towards the trichoblast zone is observed at t=15 and t=30. Red arrows show increasing root surface attachment over time (Adapted from Prieto *et al.*, 2011).

Sorghum and pearl millet tissues have a high abundance of potential siderophore-producing bacteria (SPBs). As previously discussed, siderophores are important in sequestration of iron in iron-poor environments and making it available to the bacterium and the host plant (Saha *et al*., 2012). They also have antimicrobial properties that can limit the growth of plant pathogens. SPBs are often targeted for their potential in soil bioremediation (Pilon-Smits, 2005). Siderophores bind and immobilise toxic metals, thus reducing their concentration in the soil. This is one of the mechanisms for bacterial metal tolerance (Rajkumar *et al*., 2009). Plant-associated SPBs can confer metal tolerance to their host plants, thus enabling the plant to grow in metal polluted environments (Pilon-Smits, 2005). Plant-associated metal-tolerant SPBs have been isolated from plant tissues and rhizospheric soils (Table 4.3), and these include representatives of dominant bacterial genera found in sorghum and pearl millet tissues. Therefore, these bacteria can be targeted for phytoremediation purposes for treatment of polluted agronomic soils (Pilon-Smit, 2005).

Equally important in phytoremediation technology are plant-associated bacteria that are able to degrade toxic xenobiotic compounds. Xenobiotic compounds are ecologically harmful. They are introduced into soils through industrial and agricultural practices. Bacterial species belonging to genera *Sphingomonas*, *Sphingobium*, *Exiguobacterium*, *Bacillus*, *Arthrobacter*, *Pseudomonas*, *Chryseobacterium* and *Stenotrophomonas* have been shown to degrade a wide variety of xenobiotic compounds including aromatic compounds (benzene, toulene, phenols), nitroaromatic compounds and polynuclear aromatic hydrocarbons (Radianingtyas *et*

al., 2003; Ye *et al.*, 2004; Lal *et al.*, 2008; Raina *et al.*, 2008; Krishna and Philip 2009; Chen, 2012; Jeswani and Mukherji, 2013).

Table 4.3. Examples of rhizospheric and endospheric metal tolerant siderophore producing bacteria associated with plants growing in metal contaminated environments (Adapted from Rajkumar *et al.*, 2009).

Origin of SPB	Metal contamination	Identified SPB	Reference
Rhizosphere of *Brassica juncea*	Cd	*Variovorax paradoxus*, *Flavobacterium* sp., *Rhodococcus* sp., *Ralstonia* sp., *Arthrobacter* sp., *Stenotrophomonas* sp., *Pseudomonas* sp.	Belimov *et al.*, 2005
Rhizosphere samples	Pb	*Pseudomonas* sp., *Serratia marcescens*, *Streptomyces* sp.	Kuffner *et al.*, 2008
Rhizosphere of perennial *Graminaceae* grasses	Cd, Ni, Cu	*Microbacterium* sp., *Serratia liquefaciens*, *Pseudomonas tolaasii*, *Pseudomonas fluorescens*, *Ralstonia taiwanenses*, *Agrobacterium tumefaciens*, *Paracoccus* sp., *Cellulomonas* sp.	Dell'Aminco *et al.*, 2005
Rhizosphere of *Thlaspi goesingense*	Ni	*Methylobacterium mesophilicum*, *Methylobacterium extorquens*, *Methylobacterium* sp., *Burkholderia terricola*, *Okibacterium fritillariae*, *Rhodococcus fascians*, *Rhodococcus* sp., *Microbacterium* sp.	Idris *et al.*, 2004
Shoot tissues of *Thlaspi goesingense*	Ni	*M. mesophilicum*, *Methylobacterium* sp., *Sphingomonas* sp., *Curtobacterium* sp., *Rhodococcus* sp.	Idris *et al.*, 2004
Tissues of *Allysum bertolonii*	Ni	*Staphylococcus* sp., *Microbacterium* sp., *Pseudomonas* sp., *Curtobacterium* sp., *Bacillus* sp., *Arthrobacter* sp., *Paenibacillus* sp., *Leifsonia* sp.	Barzanti *et al.*, 2007

At the time of study, there was little reason to suspect accumulation of toxic metals or complex xenobiotic compounds on the pearl millet field as the input of agrochemicals was minimal. Fertilizers applied on both plots contained nitrogen, phosphorus and potassium. The active ingredient of the herbicide, Kombat, used on the sorghum field is carbaxyl, which is a broad-spectrum insecticide that is also moderately toxic to humans and animals (EPA, 2004). This compound is considered to be easily degradable in the soil through photodegradation and microbial action. *Pseudomonas* species such as *P. putida* and *P. cepacia* are typical carbaxyl degraders (Venkateswarlu *et al.*, 1980; Chapalamadugu and Chaudhry, 1991; Chen *et al.*, 2009; Jaimini *et al.*, 2012). This could be the reason for the high abundance of *Pseudomonas* species in sorghum roots.
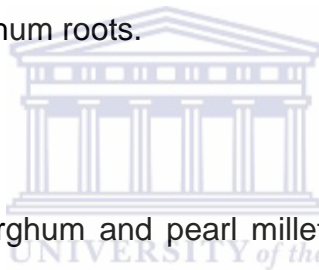
Dominant genera found in sorghum and pearl millet tissues have the potential for biocontrol applications. *Pseudomonads*, *Bacillus* and *Paenibacillus* include antibiotic-producing antipathogenic bacterial strains (Cho *et al.*, 2007; Aravind *et al.* 2009; Fürnkranz *et al.*, 2011). For example, *Pseudomonas* and *Bacillus* strains were shown to produce antibiotics that reduced growth of wilt-causing *Ralstonia solanacearum* in eggplant (Ramesh *et al.*, 2009). Strong endophytic siderophore producers like *P. aeruginosa* were also shown to suppress plant disease indirectly by outcompeting plant pathogens (e.g. *Fusarium oxysporum* and *Botrytis cinerea*) through siderophore-mediated iron sequestration (Raaijmakers *et al.*, 1995; Audenaert *et al.*, 2002). Bacteria with antipathogenic properties can be applied as live inoculum *in planta* or on agricultural soils for biocontrol of plant pathogens. An important consideration for strains targeted as biocontrol agents is that they should be fast and efficient colonisers of plant tissues (Mercado-Blanco and Bakker, 2007).

Findings in this study suggest that genera such as *Agrobacterium*, *Pseudomonas*, *Erwinia* and *Stenotrophomonas* are competent colonisers of sorghum and/or pearl millet tissues. Therefore, dominant genera with well-known biocontrol strains such as *Pseudomonas*, (Figure 4.9) can be specifically isolated and tested for their biocontrol potential in sorghum and pearl millet tissues.

Not all endophytic bacteria found in sorghum and pearl millet are beneficial. It is therefore important to note that some bacterial phylotypes found in sorghum and pearl millets could potentially be pathogenic to plants or animals, including humans. In fact, some bacterial genera known for their beneficial properties in plant-microbe interactions (e.g., *Pseudomonas*, *Agrobacterium* and *Erwinia*) are also known to consist of well-known phytopathogens (Mansfield *et al*., 2012). Even a specific bacterial species could have pathogenic and non-pathogenic strains. A case in point is the *Pseudomonas syringae* which includes naturally occurring commensalistic strains and phytotoxin-producing strains that cause blights, galls and leaf spots in a wide range of plant hosts (Bender *et al*., 1999; Mohr *et al*., 2008). These strains could have the same encoding at the 16S rRNA level; however, distinguishing virulence factors are often encoded on plasmids, pathogenicity islands, mobile elements or monocistronic genes (Pupo *et al*., 1997; Mohr *et al*., 2008; Kamar *et al*., 2013). Examples of other potential phytopathogens include soft-rot causing and fire blight inducing species in the *Erwinia* genus such as *Erwinia carotovora* and *Erwinia amylovora* (Basset *et al*., 2000; Oh and Beer, 2005). Due to their high dominance in sorghum and pearl millet tissues, these genera should receive further attention as potential pathogens of these crops.

The *Pseudomonas* genus also includes human pathogens such as the cystic fibrosis-causing *P. aeruginosa* (Govan and Deretic, 1996). Other genera consisting of human pathogens include *Rhodococcus* (e.g. *R. equi*), *Bacillus* (*B. cereus*) and *Leuconostoc* (*L. mesenteroides*) (Weinstock and Brown, 2002; Bou *et al*., 2008; Bottone, 2010). The potential presence of human pathogenic bacteria in sorghum and pearl millet has important implications for post-harvest food safety processes. This finding also raises the need for precautions to be taken regarding the application of PGPEBs on food crops. For example, a potential biofertilizer bacterium, *B. cereus*, was previously shown to produce growth inducing phytohormones (Joo *et al*., 2004); however this species is also known to cause a range of medical conditions including food poisoning, pneumonia, sepsis and central nervous system infections (Bottone, 2010). Trasmission of human pathogens via food crops has previously been shown to be possible, although this aspect has not been well-studied with regards to biofertilizer or biocontrol strains (Berger *et al*., 2010).

In light of observations discussed above, it is apparent that this study has highlighted the importance of sorghum and pearl millet plants as sources of agriculturally important bacteria. Based on these findings, future studies of these communities can include cultivation processes to isolate bacteria of specific interest. For example, diazotrophic bacteria can be targeted using nitrogen-free media such as those developed by Kirchhof *et al*. (1997). PVK and NBRIP media can be used for isolation of phosphate solubilising bacteria (Nautical *et al*., 1998). This study shows that some genera (e.g., *Pseudomonas* and *Herbaspirillum*) potentially consist of both beneficial and pathogenic species; therefore culturing would assist in determining the actual role of bacterial community members in plant hosts.

Metabolic properties of endophytic bacteria associated with sorghum and pearl millet tissues can also be exploited for many other industrial applications. Actinobacteria are an important group of bacteria that produce a broad range of biotechnologically significant metabolites such as enzymes, antimicrobial compounds, anticancer compounds, insecticides and pigments (Balagurunathan and Radhakrishnan, 2010). Bacterial antimicrobial and anticancer compounds are often used in the medical field for the treatment of human and animal diseases (Ryan *et al.*, 2007). Siderophores produced by bacteria can also be used for drug delivery whereby the siderophore is conjugated with an antibiotic so that the antibiotic is co-transported with the siderophore to the targeted location within the body. This approach is known as the Trojan-Horse strategy (Miethke and Marahiel 2007; Saha *et al.*, 2012).
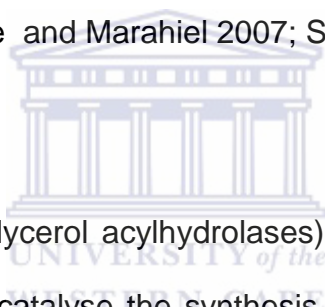
Carboxylesterases (or triacylglycerol acylhydrolases), commonly known as lipases, are a group of enzymes that catalyse the synthesis and degradation of long-chain acylglycerol (Jaeger and Eggert, 2002; Gupta *et al.*, 2004). Due to their stereoselectivity, regioselectivity and chemoselectivity, these enzymes lend themselves to a variety of biotechnological applications including bioethanol production, polymeric material biosynthesis, fine chemical production (e.g. agrochemicals, flavouring agents, cosmetics) and production of antimicrobial compounds (Jaeger and Eggert, 2002). Well-known lipase producers include strains from the *Pseudomonas*, *Bacillus*, *Burkholderia* and *Stenotrophomonas* genera (Gilbert, 1993; Jaeger and Eggert, 2002; Gupta *et al.*, 2004; Guncheva and Zhiryakova 2011; Basan-Beikdashki *et al.*, 2012).

The high abundance of genera with potential for degradation of plant polymers in sorghum and pearl millet tissues is not surprising. Competent endophytes are expected to be able to digest wood constituents (cellulose, lignin, xylan and pectin) in order to migrate to a suitable niche within the plant (Cho *et al.*, 2007). However, enzymes involved in these degradative processes also have varied biotechnological applications. Most notable are the lignocellulolytic enzymes that can be used to hydrolyse plant matter during fermentative production of biofuels. For example, an *Exiguobacterium* species has been shown to produce a range of lignocellulolytic enzymes, including cellulase, pectinase, mannanase, xylanase and tannase, during the fermentative degradation of sugarcane baggase (Vijayalaxmi *et al.*, 2013). Other bacterial genera found in sorghum and pearl millet tissues with known species that produce plant polymer degrading enzymes include *Curtobacterium*, *Pseudomonas*, *Bacillus*, and *Erwinia* (Bissaria, 1991; Lednická *et al.*, 2000).
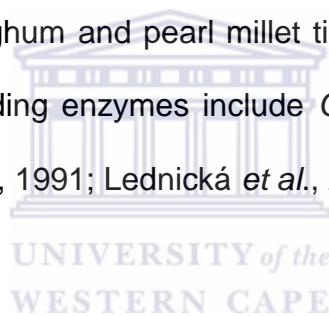
A small proportion of sequences retrieved from sorghum and pearl millet tissues could not be assigned to a taxonomic group through the CloVR pipeline. Most of these matched previously uncultured bacteria from environmental samples, with sequence similarity match of 97% or less. These sequences could belong to novel or previously uncultured bacteria. Also, bacterial phylotypes that are yet to be characterised with regards to their associations with plants were also found. These include the purple-pigmented *Janthinobacterium* and the more recently described facultative methylotrophic *Methyloversatilis* (Gillis and De Ley, 2006; Kaluyzhnaya *et al.*, 2006). Also not included in the discussions were the rare phylotypes (<1% dominance in all tissues), due to limitation imposed by lack of replicates. The rare phylotypes are highly diverse (Figure 4.6) and represent 9.5% of the overall

observed community. These considerations imply that there is great potential for discovery of novel bacterial phylotypes and/or metabolic pathways from sorghum and millet tissues, over and above the dominant groups that were discussed (Janda and Abbott, 2007).
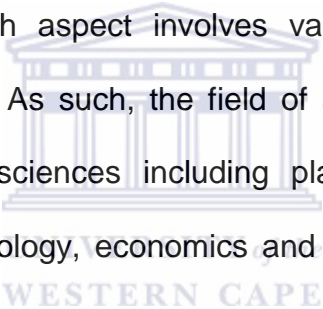
CHAPTER 5

Final Discussion

Agronomy is the branch of agriculture concerning the development of farming practices, principles and technology for production of healthy and high yielding crops (Lichtfouse *et al.*, 2009). All activities in this field are underpinned by the fundamentals of sustainable agriculture, a science that promotes responsible practices to sustain crop production, ecological soundness, commercial success and social development (Neher, 1992; Lichtfouse *et al.*, 2009). Essentially, there are three main aspects to be considered for sustainable production of healthy crops, these are physical (e.g. soil structure, climate), chemical (e.g. soil chemical composition, nutrient composition, pesticide and fertilizer use) and biological (e.g. plant disease, ecological interactions) (Schoenholtz *et al.*, 2000, Requena *et al.*, 2001; Passioura, 2007). Each aspect involves various factors that are mostly interlinked and co-dependent. As such, the field of agronomy was founded in the amalgamation of numerous sciences including plant physiology, genetics, soil science, climatology, biotechnology, economics and many more (Lichtfouse *et al.*, 2009).

The current study explores one biological influence of plant life, and that is the endophytic bacterial community found inside the plant. This study focuses on the diversity of endophytic bacterial communities associated with important African crops: sorghum, groundnut and pearl millet. The first aspect of this study was to assess a fundamental technique in molecular analysis of endophytic bacterial communities, and that is DNA extraction (Chapter 3). Here, it was shown that the choice of DNA extraction protocols (classical or commercial kits), significantly impacted the results obtained by affecting DNA yield and quality, as well as the diversity of the endophytic communities that could be detected. In the second part of

the study, high-throughput pyrosequencing was conducted to show that endophytic communities associated with sorghum and pearl millet tissues are phylogenetically and functionally diverse (Chapter 4). Bacterial endophytes in these crops can be targeted for a wide range of agricultural and industrial applications.

On the basis of findings made in this study, continued research of endophytic bacterial communities associated with sorghum and pearl millet is recommended. Effective methodological approaches are proposed to develop a better understanding of the structure of these endophytic communities and their interactions with the plant. Importantly, these communities are naturally associated with these South African crops; therefore, this study raises the need to align their exploitation with current agronomic strategies that aim to use sustainable and environmentally-friendly approaches in the production of high-yield crop and management of crop disease.

Review of methods used in the study of endophytic bacterial communities

Molecular techniques used in the study of endophytic bacterial communities are varied, as discussed in Chapter 1. Each technique has its benefits and drawbacks; therefore it is important to align the selection of the technique used with the specific research questions. Molecular techniques used in this study, t-RFLP and pyrosequencing, are both culture-independent, and were both used to analyse the distribution of the bacterial 16S rRNA gene in the metagenomic DNA from sorghum,

pearl millet and/or groundnut tissues. However, the two methods are fundamentally different in terms of their application and the type of data they generate.

Pyrosequencing provides greater resolution of bacterial communities than community profiling or culturing techniques (Marschner *et al.*, 2005). This was evident in the current study, where pyrosequencing revealed more diverse sorghum-associated communities than t-RFLP. T-RFLP retrieved up to 35 OTUs per metagenomic DNA sample, whereas pyrosequencing recovered over 300 OTUs per sample. Identification of retrieved t-RFLP OTUs by *in silico* digestion failed to discriminate between phylogenetic groups that generate similar t-RF patterns, particularly closely related species and genera. These results were in agreement with findings in the study of Ramond *et al.* (2013), where only up to 37 bacterial OTUs were detected from individual sorghum tissues using t-RFLP, although none could be assigned to one bacterial phylotype. In contrast, over 90% of OTUs retrieved by pyrosequencing from sorghum tissues were successfully assigned to genus level.

T-RFLP has already been shown to have low resolution when identifying bacterial phylotypes at a fine taxonomic scale (Blackwood *et al.*, 2007; Schütte *et al.*, 2008). Its limitations are technique-specific or PCR-related as previously discussed. However, the purpose of t-RFLP in the current study was to compare the effects of DNA extraction protocols on the diversity of microbial communities, and this only requires broad analysis of community shifts to depict these effects. Therefore, the

resolution of community diversity at this coarse taxonomic scale was sufficient to meet the intended objective in this study (Fierer, 2007).

Challenges common to all molecular techniques used in the analysis of bacterial 16S rRNA from metagenomic DNA include issues related to research design, DNA quality and PCR bias. In the current study, replication of t-RFLP analyses enabled sound statistical validation of data to determine reproducibility of results and the significance of differences observed. The lack of replicates in the pyrosequencing analysis of sorghum and pearl millet communities limited the depth of analysis that could be extrapolated (Prosser *et al.*, 2010). In order to take full advantage of the high resolution of pyrosequencing, future studies should include replicates and possibly expand to include surrounding environments in order to determine bacterial phylotypes that are specifically enriched in sorghum and pearl millet tissues (Charlson *et al.*, 2012). Previously, a geographical study conducted in South Africa using t-RFLP revealed that the sorghum rhizospheric communities were affected by abiotic factors including pH, total nitrogen and carbon content as well as clay content; however, analyses were inconclusive regarding the effects of these abiotic factors on endophytic communities (Ramond *et al.*,2013). Next generation sequencing can be used in future studies of sorghum and pearl millet endophytic communities that also include geographic and temporal elements in order to conduct an in-depth investigation of the roles that abiotic and seasonal factors play in shaping these communities. These studies would help identify core sorghum and pearl millet communities that are least affected by environmental factors.

The impact of DNA quality and yield in the study of plant associated endophytic bacteria was discussed thoroughly in Chapter 3. Therefore, this study serves as a benchmark in molecular analyses of endophytic communities associated with different crops, because it urges the evaluation of routine methodologies with respect to DNA quality and its impact on the accuracy of results. Other procedures that could affect DNA yield and quality, including plant tissue handling, storage and homogenisation processes (Varma *et al.*, 2007) as well as the handling of metagenomic DNA itself can be evaluated in future (Lahiri and Schnabel, 1993). These studies are important in the establishment of research practices that promote the quality of information generated in studies of microbial communities associated with plants and other environments.
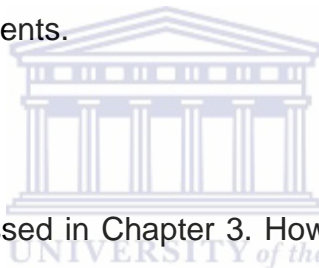
PCR biases have been discussed in Chapter 3. However, one PCR-related aspect, which is also arguably the most important, that deserves much research attention is the use of phylogenetic markers and design of primers thereof (Marschner *et al.*, 2005). At present, none of the "universal" primers targeting the bacterial 16S rRNA gene, including those used for t-RFLP and pyrosequencing tests in this study, are able to access all known bacterial phylotypes (Baker *et al.*, 2003; unpublished findings in IMBM). This means that endophytic bacterial communities studied via analysis of this phylogenetic marker are potentially underestimated. Primer design is continuously researched, with intentions of developing "universal" primers that are target-specific, and yet able to accurately capture the diversity of native communities (Marchesi *et al.*, 1998; Baker *et al.*, 2003).

The length of pyrosequencing primers can introduce bias to bacterial community studies. Multiplexed barcoded primers, as used in this study, are long (60bp). The barcode and the adaptor sequences on the primer are not gene-specific, and these introduce bias in the amplification of the bacterial 16S rRNA gene. Primers with different barcodes (i.e., barcode sequences are different), when used in one experiment, introduce uneven bias in the experiment (Berry *et al.*, 2011). Also, long primers require a high annealing temperature (Wu *et al.*, 1991) – 72°C in this study – but this increases primer bias towards certain phylotypes in mixed-template PCR reactions (Hongoh *et al.*, 2003). To mitigate these problems, a 2-step PCR reaction is proposed for future analysis of crop-associated bacteria. In this case the target sequence is first amplified from the metagenomic DNA using target-specific primers at a low annealing temperature, followed by amplification of the target gene from the generated amplicons using multiplexed primers (Berry *et al.*, 2011). The first amplification liberates the target genes from the metagenomic DNA pool, thus increasing their accessibility to the multiplexed primer in the second round of PCR.

Another pyrosequencing-specific problem is the limited sequence read-length that pyrosequencing platforms can generate. The Roche GS Junior pyrosequencer used in this study can achieve read-lengths of 400-500bp. The actual maximum length of sequences generated was 470bp. These reads are much shorter than those generated by traditional Sanger sequencing technology (>1000bp) (Chan, 2005). Short sequences limit the resolution of bacterial taxonomies, particularly at species or strain level; and hence, identification of bacterial phylotypes was restricted to genus level (Janda and Abbott, 2007). All next generation sequencing platforms aim

to increase read length as they continue to improve, in order to allow for greater sequence coverage and increased sensitivity (Okubo *et al.*, 2012).

The efficiency of the 16S rRNA gene sequence as a phylogenetic marker is often questioned. This gene has very low resolution power at fine taxonomic levels (i.e. species or strain) and provides little information on the functional properties of community phylotypes and their relatedness (Janda and Abbott, 2007; Petrosino *et al.*, 2009). Santos and Ochman (2004) have previously proposed the use of multi-gene specific primers sets to target protein loci sequences for phylogenetic community characterisation in order to resolve the phylogenetic classifications and evolutionary traits of communities. Another alternative could be to target and analyse the distribution of specific functional gene sequences such as *nifH* (nitrogenase activity) (Zehr *et al.*, 2003) and *amoA* (ammonium oxidation) (Rotthauwe *et al.*, 1997) as phylogenetic markers in order to focus the study on the diversity of bacteria with specified metabolic properties. However, even though both approaches provide solutions to some of the problematic areas, they are still encumbered by the limitations of single gene analysis.

To date, whole metagenome shotgun sequencing is the most holistic approach in the study of environmental microbial communities (Petrosino *et al.*, 2005). This technique bypasses the single-gene analysis and its associated PCR and priming challenges because the metagenomic DNA is fragmented and sequenced directly thereafter. This means that all genes, phylogenetic and functional, are sampled in this approach, thus allowing for a comprehensive analysis of all the possible

ecological interactions taking place in the environment (Petrosino *et al.*, 2005). For example, pyrosequencing of 16S rRNA and 18S rRNA genes conducted in parallel to whole-genome shotgun pyrosequencing was recently used to study the structure of bacterial and fungal endophytes associated with different tissues of a tomato plant, and to identify elements within the plant that are related to its susceptibility to a pathogenic *Salmonella* infection (Ottesen *et al.*, 2013). This method can be used in future studies of sorghum and pearl millet communities because the current study does not delve into specific interactions taking place between the plants and their associated microbiome. Other important pyrosequencing-based techniques are metatranscriptomics and metabolomics, which can be used to study gene expression and metabolite producing patterns in the endosphere, in order to further elucidate important metabolic activities taking place (Bundy *et al.*, 2009; Gilbert and Hughes, 2011).
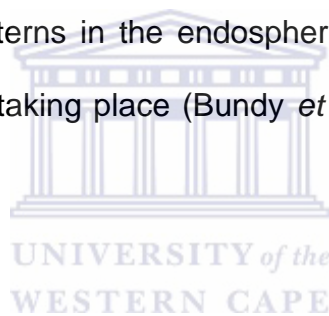
The current study revealed taxonomically and metabolically diverse bacterial communities in tissues of sorghum and pearl millet. Therefore, bacteria of specific phylogenetic affiliation or functional characteristics can also be specifically targeted in culture-based approaches, paired with metagenomic analyses in order to further characterise them in terms of their genotypic and phenotypic properties, as well as their specific role(s) in the endosphere. An example would be the integrated study conducted by Weston *et al.*, (2012), that elucidates the interaction of *Pseudomonas fluorescens* GM30 and *P. fluorescens* Pf-5 strains (both native to *Populus deltoides*) with *Arabidoposis thaliana*. These strains were first shown to have growth promoting properties (IAA and siderophore production) in culture-based studies, and shown to increase lateral root biomass in *in vivo* inoculation studies. Furthermore, a

metatrascriptomics analysis of the plant genome was used to survey changes in the plant's gene expression pathways when inoculated with these beneficial strains (individually or together) in the presence or absence of a pathogenic strain, *P. syringae* DC3000. Whole genome sequencing of individual bacterial cells can also be conducted on important strain, in order to analyse an entire array of their genes and their potential activities in the plant (Ng and Kirkness, 2010).

Prospects of sorghum and pearl millet endophytome in South Africa

Potential for biofertilization and biocontrol development:

Biotechnological innovations have sought to address crop and land management issues in order to sustain productivity of agronomic land. One such approach is the use of plant-associated microorganisms for plant-growth promotion (biofertilizers) and disease management (biocontrol).

The current study confirms that local crops, sorghum and pearl millet, are naturally associated with bacterial genera that have biofertilizer or biocontrol potential. Some of these, including *Curtobacterium*, *Microbacterium*, *Rhodococcus*, *Chryseobacterium*, *Bacillus*, *Paenibacillus*, *Exiguobacterium*, *Sphingobium*, *Sphingomonas*, *Herbaspirillum*, *Erwinia*, *Pseudomonas* and *Stenotrophomonas*, are dominant genera in both sorghum and pearl millet tissues (Chapter 4), and were previously isolated in other agricultural crops (Pereira *et al*., 2011; Magnani *et al*., 2013). Other genera with plant growth promoting properties appeared to be plant

tissue-specific in this study; for example, *Arthrobacter* and *Exiguobacterium* were only found in the millet root, and *Rhizobium* were only found in sorghum stem and root. However, these genera have been observed in other grass crops such as rice, maize and wheat in previous studies (James and Olivares, 1997; Franche *et al.*, 2009; Pereira *et al.*, 2011; Pisarka and Pietr 2012). Therefore, potential PGPEBs identified in sorghum and pearl millet tissues are adaptable to a broad range of host plants. This is an ideal property of a broad-application biofertilizer or biocontrol agent (Mohammadi and Sohrabi, 2012).

Another attractive quality of bacterial genera in sorghum and pearl millet is their potential for multiple plant growth promoting activities (Figure 4.7). Genera such as *Bacillus* and *Pseudomonas* are particularly known for their broad range of plant-growth inducing metabolic capabilities as discussed in Chapter 4. Also, dominant genera such as *Pseudomonas*, *Erwinia* and *Stenotrophomonas*, were represented by over ten OTUs in both plants. This could be a reflection of the diversity of their representative species and strains in these tissues, which would emphasize the importance of these metabolically diverse genera in sorghum and pearl millet.

PGPEBs can also be tested for their ability to alleviate environmental stress in crops, particularly in the Sub-Saharan regions where droughts, elevated irradiation and desertification are prevalent. Sorghum and pearl millet are relatively drought-tolerant (Belton and Taylor 2004); therefore, they are plausible sources of bacteria that might be involved in conferring drought-tolerance in plants. Indeed, genera such as *Pseudomonas* and *Bacillus*, which are dominant in these plants, were previously

implicated in this role. *P. putida* and *B. megaterium* were shown to improve osmotic stress tolerance in *Arabidopsis thaliana* through increased production of proline and IAA (Marulanda *et al.*, 2009). In that study, co-inoculation of *B. megaterium* with arbuscular mycorrhizal fungi further enhanced the plant's water stress tolerance. A similar observation was made in a previous study when *B. thuringiensis* was co-inoculated with three mycorrhizal fungi species in the legume, *Retama sphaerocarpa* (Marulanda *et al.*, 2006). These bacteria, when isolated from sorghum and pearl millet tissues, can be assessed in other important South African crops that are more susceptible to drought such as maize (Castiglioni *et al.*, 2008). The role of other microorganisms (e.g. fungi) associated with sorghum and pearl millet should also be explored.

The first step in the development of biofertilizers or biocontrol strains from the sorghum and pearl millet endophytome would be to isolate the key bacteria identified in this study which potentially produce specified properties (Mohammadi and Sohrabi, 2012). General culturing media can be used to target all culturable bacteria; whereas selective and semi-selective media can be used to isolate bacteria with specific properties. Isolated bacteria can be characterised based on predefined properties, and strains with enhanced plant growth promoting properties would be used in *in vivo* inoculation trials, to assess their specific physiological effects on the host plant. PGPEBs can be tested individually or in consortia, whereby the latter option would typically include several strains that confer different benefits to the plant. The advantage of developing PGPBEs isolated directly from indigenous crops for bio-inoculation processes is that these bacteria are already well-adapted to the endospheric environment (Hardoim *et al.*, 2008); therefore, they would require little

or no genetic modification prior to development of inocula for agricultural applications. Of course, not all endophytic bacteria identified through metagenomic analyses are readily culturable in laboratory; however, microbiological techniques are growing more elaborate and creative in order to access more bacteria from environmental samples (Vartoukian *et al.*, 2010; Stewart, 2012).

There are commercial producers of biofertilizer and biocontrol products in South Africa already (Barnard and du Preez, 2004). However, there is an acute lack of information regarding the large-scale or long-term impact of these products in Sub-Saharan Africa (Barnard and du Preez, 2004; Chianu *et al.*, 2010). Most publications on the subject report only on *in vivo* or small-scale field experiments. The lack of resources to expand these studies is blamed for this situation (Chianu *et al.*, 2009). However, South Africa, as the economic leader in this region, is in a position to expand research and make it commercially viable for itself and its neighbouring countries.

PGPEBs isolated from sorghum and pearl millet can also be evaluated in the context of other South African farming practices and technologies aimed at improving plant yield and/or health. As agricultural processes become more integrated, this focus would position the use of PGPEBs well within future agronomic strategies (Tilak *et al.*, 2005; Wu *et al.*, 2007). One such consideration would be the assessment of interactions between PGPEBs and genetically modified (GM) crops. African countries are increasingly farming high-yield genetically modified crops (including sorghum and millets) which are tolerant to environmental stresses and/or disease

resistant (Cohen, 2005). Enhanced properties of GM crops could affect colonisation patterns of beneficial microorganisms. For example, it was previously shown that increased production of antibacterial metabolites in GM crops affects the structure of their associated native endophytic communities (Rasche *et al.*, 2006). Therefore, colonisation of selected PGPEBs can be analysed in sorghum and pearl millet varieties farmed in South Africa.

Endophytic bacteria can also be used in the development of transgenic crops. In one approach, bacterial genes can be transformed into crop plants to produce healthy and fit transgenic varieties. In fact, the first GM crop farmed in South Africa was the Bt maize (Gouse *et al.*, 2005), which is maize crop transformed with *Cry* genes from *Bacillus thuringiensis* for production of Cry proteins that are toxic against a range of insect pests (Hellmich and Hellmich, 2012). In another approach, endophytic bacteria such as pathogenic *Agrobacterium* species (*A. tumefaciens* and *A. rhizogenes*) are used in horizontal gene transfer (HGT) processes, for mobilisation of genetic material into plant cells during the development of transgenic crops (Chilton *et al.*, 1982). In this process, the Ti or Ri plasmid of the bacterium that carries tumor-inducing genes is "disarmed" (i.e., virulence genes are deactivated), and the desired gene sequence is ligated into the plasmid, which is then transferred into the plant via the natural *Agrobacterium*-mediated transformation process. This process is routinely used in the transformation of important agricultural crops including maize, sorghum, pearl millet and wheat (Risacher *et al.*, 2009; Saikishore *et al.*, 2011; Ramadevi *et al.*, 2014). However, the transformation success rate is still very low due to the instability of the plasmid within the host plant and the plant's defense system against the invading microbial particles (Saikishore *et al.*, 2011; Pitzschke *et*

al., 2013). Therefore, more efficient strains are continuously developed to improve this process for transformation of agricultural crops (Aarrouf *et al*., 2011; Jha *et al*., 2011; Wu *et al*., 2014). Other genera such as the *Rhizobium* are also considered for their HGT capabilities (He *et al*., 2003; Broothaerts *et al*., 2005). These genera, *Agrobacterium* and *Rhizobium*, were abundant in sorghum tissues in the current study; therefore future studies can also target these bacteria to test them for their HGT efficiency.

Bioremediation of contaminated land:

Sorghum and pearl millet tissues have a high abundance of bacteria with potential bioremediation properties. These include bacteria that are capable of breaking down complex xenobiotic compounds and those with high metal tolerance. Essentially, all dominant genera found in sorghum and pearl millet tissues, except for *Leuconostoc*, *Swaminathania*, *Methyloversatilis* and *Janthinobacteria*, have at least one species capable of either one of these properties (Figure 4.7). Of these, genera such as *Pseudomonas*, *Erwinia*, *Stenotrophomonas*, *Agrobacterium* and *Microbacterium* represent the most dominant groups that were found in both plants. Interestingly, Sphingomonads, *Sphingobium* and *Sphingomonas*, which are well-known for their broad biodegradation properties (Stolz, 2009) were only abundant in sorghum root and stem, respectively.

Bacteria isolated from sorghum and pearl millet tissues can be tested for their metal tolerance range and threshold, as well as their ability to degrade specific

compounds. Strains with enhanced metal tolerance or biodegradation capabilities can be considered for bioremediation processes. In direct bioremediation, bacteria are inoculated into the soils where they can break down complex compounds (McGuinness and Dowling, 2009). In phytoremediation processes, these bacteria can be inoculated into plants, where they sequester heavy metals and/or xenobiotic compounds (Pilon-Smits, 2005). These plants, known as hyperaccumulators, are then able to extract substantial amounts of contaminants from the soil, and can later be harvested to remove the chemicals from the environment. Naturally, endophytic bacteria are suitable candidates for this process. Table 5.1. shows examples of endophytic bacteria that have previously been used in bioremediation of contaminated soils, some of which belong to genera (*Herbaspirillum* and *Pseudomonas*) observed in sorghum and pearl millet tissues in the current study. Therefore, this study highlights sorghum and pearl millet as potential phytoextractors of a broad range of metals and xenobiotic compounds.
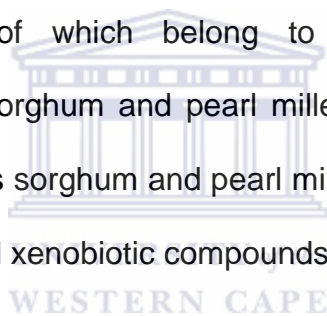
Sorghum and pearl millet crops are particularly suitable for bioremediation processes because they grow fast and have high biomass (Vamerali *et al*., 2010). Indeed, both plants were previously shown to be moderate hyperaccumulators, with the former used in phytoextraction of metals including cadmium, zinc and chromium (Epelde *et al*., 2009; Revathi *et al*., 2011), and the latter, cadmium and lead (Wuana *et al*., 2013). Enhanced metal accumulation also has a profitable application in phytomining, whereby, the same phytoextraction principles are used to recover valuable minerals such as gold and lead from soils (Wilson-Corral *et al*., 2010; Sheoran *et al*., 2013). Sorghum has already shown little potential as a gold

phytoextractor (Piccinin *et al.*, 2007); however, further research can be conducted to test for potential enhancement of this property by its endophytic bacterial community.

Table 5.1. Reported cases of successful bioremediation using endophytic bacteria (adapted from McGuinness and Dowling, 2009)

Compound	Plants used	Microbes used	Reference
PCBs, TCP	Wheat (*Triticum*, spp.)	*Herbaspirillum* sp K1	Mannisto *et al.*, 2001
Chlorobenzoic acids	Wild rye (*Elysum dauricus*)	*Pseudomonas aeruginosa* R75 *Pseudomonas savastanoi* CB35	Siciliano *et al.*, 1998
Pesticide: 2,4-D	Pea (*Pisum sativum*)	*Pseudomonas putida* VM1450	Gerrmaine *et al.*, 2009
Toulene	Yellow lupine (*Lupinus luteus* L.)	*Burkholderia cepacia* G4	Barac *et al.*, 2004
	Poplar (*Populus*)	*Burkholderia cepacia* Bu61	Taghavi *et al.*, 2005
VOCs: MTBE, BTEX, TCE	Poplar (*Populus* cv. *Hazendans* and cv. *Hoogvorst.*	*Pseudomonas* sp.	Germaine *et al.*, 2006 Porteus-Moore *et al.*, 2006
HCs: Napthalene	Pea (*Pisum sativum*)	*Pseudomonas putida* VM1441 (pNAH7)	Germaine *et al.*, 2009
Explosives: TNT, RDX, HMX	Poplar tissues (*Populus deltoidesnigra* DN34)	*Methylobacterium populi* BJ001	Van Aken *et al.*, 2004

Bioremediation technology is very relevant in South Africa where contamination of soils and water systems is high, notably due to agricultural, mining and other industrial activities. Mining and farming are arguably two of the leading causes of pollution due to the release of mine tailings into the environment and use of agrochemicals on commercial farms (Rösner and van Schalkwyk, 2000; Schulz, 2001; Reinecke and Reinecke, 2007). Traditional soil cleaning methods such as soil

excavations, washing and burning processes are costly as they require high energy input and cause further perturbation of soil environments (Scholz and Schnabel, 2006). Bioremediation processes are a cheaper alternative, that also promote recovery of natural ecosystems (Requena *et al.*, 2007).

Bioproduct production potential:

The potential of sorghum and pearl millet endophytic communities in producing industrial metabolites was highlighted in Chapter 4. Therefore, this study proposes further bioprospecting studies for specific isolation of important strains, or biomining of genes expressing these metabolites. These events could include culturing and metagenomic processes. With the former approach, bacteria can be isolated from plant tissues and screened for specific activities such as lipase and antibiotic production (Guncheva and Zhiryakova 2011; Vijayalaxmi *et al.*, 2013), and these activities can be optimised to develop feasible large-scale industrial application. The disadvantages with this approach are the associated high cost and intensive labour requirements (Taylor *et al.*, 2012). In metagenomic approaches, environmental (plant tissue) metagenomic DNA can be fragmented and cloned into vector for expression in vehicle organisms; and gene expression would then be screened against a background of relevant activities (Taylor *et al.*, 2012). The latter option is particularly attractive in targeting activities of unculturable obligate endophytes that cannot grow in culture, which generally represent 99% of the organisms (Aslam *et al.*, 2010).

Targeted production of bioproducts from sorghum and pearl millet tissues would be in line with the South Africa's strategic pursuit of bio-based industries. The country adopted the National Biotechnology Plan in 2001 to stimulate the growth of the local biotechnological industry, in order to affirm its position as the African leader in providing biotechnological solutions and increase its competitiveness in global markets (Motari *et al.*, 2004; Webster and Akanbi, 2005).

Socio-economical and ecological impact:

Exploration of the sorghum and pearl millet endophytome in this study was primarily aimed at the discovery of PGPEBs for development of biofertilizers and biocontrol products to promote yield and health of these plants, and other related staple crops. Therefore, this objective directly tackles the food shortage crisis in African countries. Alleviation of poverty has far-reaching implications, which include increasing productivity of the population, and improved management of diseases such as HIV/AIDS and tuberculosis (Stige *et al.*, 2006).

Production and use of agricultural bioinoculants has potential economical value, as these can be traded to generate revenue. Use of these products can reduce use and cost of main costs of chemical fertilizers, which are largely imported (Chianu *et al.*, 2009). Increased food production and reduced farming costs can lead to lower food prices.

The emerging bio-based industries are also a source of job creation and for the poor in the country. Obviously, developing biotechnological companies would employ staff, and outsource supporting services. Also, activities surrounding the biotechnological developments can stimulate entrepreneurship. An example of this is the outsourcing of hyperaccumulator plant supply to the smallholder businesses or community-based farms (AngloGold Ashanti, 2004). Application of biofertilizers is uncomplicated, and as it was with the roll-out of GM foods (Gouse *et al.*, 2005), these products can be made available to subsistence farmers to increase food security in rural areas.

Research on crop associated bacteria and their applications offers environmentally-friendly strategies that can be applied in varied industries. Such applications, already discussed in this chapter, include the use of biofertilizers and biocontrol agents as alternatives to toxic and ecologically-harmful agrochemicals and phytoremediation processes that alleviate the soil of toxic chemicals. These processes also promote establishment of key ecological pathways; thus preserving biodiversity in the midst of increasing urbanisation and industrialisation (Requena *et al.*, 2001).

CONCLUSIONS AND PERSPECTIVES

In summary, this study has shown that previously underexplored sorghum and pearl millet endophytic bacterial communities are highly diverse. They consist of bacterial groups with potential for plant growth promotion and a broad range of metabolic capabilities that can be exploited for industrial applications. Dominant bacterial genera such as *Pseudomonas*, *Paenibacillus*, *Bacillus*, *Agrobacterium*, and *Erwinia* can be targeted for development of broad spectrum biofertilizers and biocontrol agents, which can be inoculated into these crops to increase crop yield and manage plant diseases. Genera such as *Herbaspirillum*, *Sphingobium*, *Sphingomonas*, *Swaminathania* and *Rhizobium* were only dominant in sorghum tissues; and *Arthrobacter*, *Chryseobacterium*, *Janthinobacterium* and *Methyloversatilis* were most dominant in pearl millet tissues. It is possible that these bacteria are specifically enriched in these plants; however, a broader study is required to cofirm this. Subsequent studies could integrate culturing, community profiling techniques and high-resolution next generation sequencing techniques (e.g. whole metagenome shotgun sequencing, metatrascriptomics) to include assessment of surrounding communities (phyllospheric and rhizospheric), monitor seasonal variations, assess environmental stress impact and elucidate interactions between the plant and important strains.

Routine procedures (e.g. sample handling, DNA extraction, PCR, etc.) that are involved in the study of environmental bacterial communities could have an effect on the observed community structure and diversity. The current study shows that DNA

extraction protocols introduce a bias in the diversity of endophytic bacterial communities. The efficiency of different protocols is affected by the plant species and the quantity of tissue used. Cell lysis and DNA purification steps included in a DNA extraction procedure are considered to be the most crucial steps in the retrieval of diverse phylotypes from the plant tissues. Therefore, evaluation of these procedures should be considered for different plant matrices.

REFERENCES

Aarrouf J., Castro-Quezada P., Mallard S., Caromel B., Lizzi Y. and Lefebvre V. 2011. *Agrobacterium* rhizogenes-dependent production of transformed roots from foliar explants of pepper (*Capsicum annuum*): a new and efficient tool for functional analysis of genes. *Plant Cell Reports*. **31:** 391-401.

Agu R.C. and Palmer G.H. 1998. A reassessment of sorghum for lager-beer brewing. *Bioresources Technology*. **66:** 253-261.

Ai Y., Medic J., Jiang H., Wang D. and Jane J. 2011. Starch characterization and ethanol production of sorghum. *Journal of Agricultural and Food Chemistry*. **59:** 7385-7392.

Ali Sk.Z., Sandhya V., Grover M., Kishore N., Rao L.V. and Venkateswarlu B. 2009. *Pseudomonas* sp. strain AKM-P6 enhances tolerance of sorghum seedlings to elevated temperatures. *Biology and Fertility of Soils*. **46:** 45-55.

Ambrosini A., Beneduzi A., Stefanski T., Pinheiro G., Vargas L.K. and Passaglia L.M.P. 2012. Screening of plant growth promoting rhizobacteria isolated from sunflower (*Helianthus annuus* L.). *Plant Soil*. **356:** 245-264.

Andrews M., Hodge S. and Raven J.A. 2010. Positive plant microbial interactions. *Annals of Applied Biology*. **157:** 317-320.

Angiuoli S.V.,Matalka M., Gussman G., Galens K., Vangala M., Riley D.R., Arze C., White J.R., White O. and Fricke W.F. 2011. CloVR: A virtual machine for automated and portable sequence analysis from the desktop using cloud computing. *BMC Bioinformatics*. **12:** 356.

AngloGold Ashanti Limited. 2004. Report to Society - Environment Case Study (South Africa): Woodlands project – good progress being made with phytoremediationproject. Website: http://www.anglogold.co.za. Last viewed: 18 February 2014.

Aravind R., Kumar A., Eapen S.J. and Ramana K.V. 2009. Endophytic bacterial flora in root and stem tissues of black pepper (*Piper nigrum* L.) genotype: isolation, identification and evaluation against *Phytophthora capsici*. *Letters in Applied Microbiology*. **48:** 58-64.

Aslam Z., Yasir M., Khaliq A., Matsui K. And Chung Y.R. 2010. Too much bacteria still unculturable. *Crop and Environment*. **1:** 59-60.

Audenaert K., Pattery T., Cornelis P. and Höfte M. 2002. Induction of systemic resistance to *Botrytis cinerea* in tomato by *Pseudomonas aeruginosa* 7NSK2: role of salicylic acid, pyochelin, and pyocyanin. *Molecular Plant-Microbe Interactions*. **15:**1147–1156.

Baker G.C., Smith J.J. and Cowan D.A. 2003. Review and re-analysis of domain-specific 16S primers. *Journal of Microbiological Methods*. **55:** 541-555.

Barnard R.O. and du Preez C.C. 2004. Soil fertility in South Africa: the last twenty-five years. *South African Journal of Plant and Soil*. **21:** 301-315.

Balagurunathan R. and Radhakrishnan M. 2010. Biotechnological, genetic engineering and nanotechnological potential of actinomycetes. *In* Maheshwari D.K., Dubey R.C. and Saravanamurthu R. (editors). Industrial Exploitation of Microorganisms. International Publishing House, New Delhi. Chapter 15. Pages 302-309.

Baldani V.L.D., Baldani J. I. and Döbereiner J. 2000. Inoculation of rice plants with the endophytic diazotrophs *Herbaspirillum seropedicae* and *Burkholderia* spp. *Biology and Fertility of Soils*. **30:** 485-491.

Basset A., Khush R.S., Braun A., Gardan L., Boccard F., Hoffmann J.A. and Lemaitre B. 2000. The phytopathogenic bacteria *Erwinia carotovora* infects Drosophila and activates an immune response. *Proceedings of the national Academy of Sciences of the United States of America*. **97:** 3376-3381.

Bedini S., Pellegrino E., Avio L., Pellegrini S., Bazzoffi P., Argese E., Giovannetti M. 2009. Changes in soil aggregation and glomalin related soil protein content as affected by the arbuscular mycorrhizal fungal species *Glomus mosseae* and *Glomus intraradices*. *Soil Biology and Biochemistry*. **41**: 1491–1496.

Belton P.S. and Taylor J.R.N. 2004. Sorghum and millet: protein sources for Africa. *Trends in Food Science & Technology*. **15:** 94-98.

Bender C.L., Alarcón-Chaidez F. and Gross D. 1999. *Pseudomonas syringae* phytotoxins: mode of action, regulation, biosynthesis by peptide and polyketide synthetases. *Microbiology and Molecular Biology Reviews*. **63:** 266-292.

Berg G. and Smalla K. 2009. Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. *FEMS Microbiology Ecology*. **68:** 1-13.

Berger C.N., Sodha S.V., Shaw R.K., Griffin P.M., Pink D., Hand P. and Frankel G. 2010. Fresh fruit and vegetables as vehicles for the transmission of human pathogens. *Environmental Microbiology*. **12:** 2385-2397.

Berlec A. 2012. Novel techniques and findings in the study of plant microbiota: search for plant probiotics. *Plant Science*. **193:** 96-102.

Berry D., Mahfoudh K.B., Wagner M. and Loy A. 2011. Barcoded primers used in multiplex amplicon pyrosequencing bias amplification. *Applied Environmental Microbiology*. **77:** 7846.

Bertalan M., Albano R., de Pádua V., Rouws L. Rojas C., Hemerly A., Teixeira K., Schwab S., Araujo J., Oliveira A., *et al*. 2009. Complete genome sequence of the sugar-cane nitrogen-fixing endophyte *Gluconacetobacter diazotrophicus* Pal5. *BMC Genomics*. **10:** 450.

Berthelet M., Whyte L.G. and Greer C.W. 1996. Rapid, direct extraction of DNA from soils for PCR analysis using polyvinylpolypyrrolidone spin columns. *FEMS Microbiology Letters*. **138:** 17-22.

Bhattacharjee R.B., Singh A. and Mukhopadhyay S. N. 2008. Use of nitrogen-fixing bacteria as biofertiliser for non-legumes: prospects and challenges. *Applied Microbiology and Biotechnology*. **80:** 199-209.

Bissaria, V. S. 1991. Bioprocessing of agro-residues to glucose and chemicals. *In* Martin M.A. (editor). Bioconversion of waste materials to industrial products. Elsevier. London. Pages 210–213.

Blackwood C., Hudleston D., Zak D. and Buyer, J. 2007. Interpreting ecological diversity indices applied to terminal restriction fragment length polymorphism data: insights from simulated microbial communities. *Applied and Environmental Microbiology*. **73:** 5276- 5283.

Blanco Y., Blanch M., Pin D., Legaz M.E. and Vicente C. 2005. Antagonism of *Gluconacetobacter diazotrophicus* (a sugarcane endosymbiont) against *Xanthomonas albilineans* (pathogen) studied in alginate immobilized sugarcane stalk tissues. *Journal of Bioscience and Bioengineering*. **4:** 366-371.

Bottone E.J. 2010. *Bacillus cereus*, a volatile human pathogen. *Clinical Microbiology Reviews.* **23:** 382-398.

Bou G., Saleta J.L., Nieto J.A.S., Tomas M., Valdezate S., Sousa D., Luiero F., Villanueva R., Pereira M.J. and Llinares P. 2008. Nosocomial outbreaks caused by *Leuconostoc mesenteroides* subs. mesenteroides. *Emerging Infectious Diseases.* 14. Available from http://wwwnc.cdc.gov/eid/article/14/6/07-0581.htm. Accessed 12 April 2014.

Bray J. and Curtis J. 1957. An ordination of the upland forest communities of Southern Wisconsin. *Ecological Monographs*. **27:** 325-349.

Broothaerts W., Mitchell H.J., Weir B., Kaines S., Smith L.M.A., Yang W., Mayer J.A., Roa-Rodriguez C. and Jefferson R.A. 2005. Gene transfer to plants by diverse species of bacteria. *Nature.* **433:** 629-633.

Buckley D.H. and Schmidt T.M. 2001. The structure of microbial communities in soil and the lasting impact of cultivation. *Microbial Ecology.* **42:** 11-21.

Buée M., De Boer W., Martin F., van Overbeek L. and Jurkevitch E. 2009. The rhizospheric zoo: An overview of plant-associated communities of microorganisms, including phages, bacteria, archaea, and fungi, and some of their structuring factors. *Plant Soil.* **321:** 189-212.

Bundy J.G., Davey M.P. and Viant M.R. 2009. Environmental metabolomics: a critical review and future perspectives. *Metabolomics.* **5:** 3-21.

Butt A., Mousley C., Morris K., Beynon J., Can C., *et al.* 1998. Differential expression of a senescence-enhanced metallothione in gene in Arabidopsis in response to isolates of *Peronospora parasitica* and *Pseudomonas syringae. Plant Journal.* **16:** 209-221.

Caporaso J.G., Kuczynski J., Stombaugh J., Bittinger K., Bushman F.D., Costello E.K., Fierer N., Pena A-G, Goodrich J.K., Gordon J.I. *et al.* 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods.* **7:** 335-336.

Caporaso J.G., Bittinger K., Bushman F.D., DeSantis T.Z., Andersen G.L. and Knight R. 2010. PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics.* **26:** 266-267.

Castiglioni P., Warner D., Bensen R.J., Anstrom D.C.,Harrison J., Stoecker M., Abad M., Kumar G., Salvador S., D'Ordine R. *et al.*, 2008. Bacterial RNA chaperones confer abiotic stress tolerance in plants and improved grain yield in maize under water-limited conditions. *Plant Physiology.* **147:** 446-455.

Chaiharn M., Chunhaleuchanon S. and Lumyong S. 2009. Screening siderophore producing bacteria as potential biological control agents for fungal rice pathogens in Thailand. *World Journal of Microbiology and Biotechnology.* **25:** 1919-1928.

Chan E.Y. 2005. Advances in sequencing technology. *Mutation Research.* **573:** 13-40.

Chapalamadugu S. and Chaudhry G.R. 1991. Hydrolysis of carbaxyl by a *Pseudomonas* sp. and construction of a microbial consortium that completely metabolizes carbaxyl. *Applied and Environmental Microbiology*. **57:** 744-780.

Charles T.C. 2010. The potential for investigation of plant-microbe interactions using metagenomics methods. *In* Marco D. (editor) Metagenomics: Theory, Methods and Application. Caister Academic Press. Norfolk, UK. Chapter 6, pages 107-115.

Charlson E.S., Bittinger K., Chen J., Diamond J.M., Li H., Collman R.G. and Bushman F.D. 2012. Assessing bacterial populations in the lung by replicate analysis of samples from the upper and lower respiratory tracts. *PLoS ONE.* **7:** e42786.

Chaudhary D.R., Saxena J., Lorenz N., Dick L.K. and Dick R.P. 2012. Microbial profiles of rhizosphere and bulk soil microbial communities of biofuel crops switchgrass (*Panicum virgatum* L.) and jatropha (*Jatropha curcas* L.). *Applied and Environmental Soil Science*. Vol 2012, Article ID 906864.

Chen H., He X., Rong X., Chen W., Cai P., Liang W., Li S. and Huang Q. 2009. Adsorption and biodegradation of carbaxyl on montmorillonite, kaolinite and goethite. *Applied Clay Science.* **46:** 102-108.

Chen R.S., Wu P.L. and Chiou R.Y.Y. 2002. Peanut as a source of resveratrol. *Journal of Agricultural and Food Chemistry*. **50:** 1665-1667.

Chen Q., Wei S., Deng Z., Yin L., He B. and Kong X. 2009. Optimisation of DNA extraction from seeds of *Sorghum sudanense* (Piper) Stapf. *Notulae Botanicae Agrobotanici Cluj-Napoca*. **37:** 256-260.

Chen W. 2012. The study of bioremediation on heavy metal of cultured seawater by *Sphingomonas* sp. XJ2 immobilized *Sphingomonas* strain. *Advanced Materials Research*. **347-353:** 1436-1441.

Chianu J.N., Nkonya E.M., Mairaru F.S., Chianu J.N. and Akinnifesi F.K. 2009. Biological nitrogen fixation and socioeconomic factors for legume production in Sub-Saharan Africa: a review. *Agronomy for Sustainable Development*. **31:** 139-154.

Chilton M., De Framond A., Byrne M., Fraley R., Chilton W.S., Fenning L., Barton K.A., Binns A.N., Matzke A.J.M., Bevan M., *et al*. 1982. T-DNA of the *Agrobacterium* Ti and Ri plasmids as vectors. *Structure and Function of Plant Genomes*. **63:** 409-418.

Chin-A-Woeng T.F.C., de Priester W., van der Bij A.J. and Lugtenberg B.J.J. 1997. Description of the colonization of a gnotobiotic tomato rhizosphere by *Pseudomonas fluorescens* biocontrol strain WCS365, using scanning electron microscopy. *MPMI*. **10:** 79-86.

Chisholm S.T., Coaker G., Day B. and Staskawicz B.J. 2006. Host-microbe interactions: shaping the evolution of the plant immune response. *Cell*. **124:** 803-814.

Cho K.M., Hong S.Y., Lee S.M., Kim Y.H., Kahng G.G., Lim Y.P., Kim H. and Yun H.D. 2007. Endophytic bacterial communities in ginseng and their antifungal activity against pathogens. *Microbial Ecology*. **54:** 341-351.

Christin P., Salamin N., Kellogg E.A., Vicentini A. and Besnard G. 2009. Integrating phylogeny into studies of C4 variation in the grasses. *Plant Physiology*. **149:** 82-87.

Clarke K.R. and Warwick R.M. 2nd Edition. 2001. Change in communities: an approach to statistical analysis and interpretation. Primer-E Plymouth. Chapters 1-17. Australia.

Cohen A.C., Travaglia C.N., Bottini R. and Piccoli P.N. 2009. Participation of abscisic acid and gibberellins produced by endophytic *Azospirillum* in the alleviation of drought effects in maize. *Botany*. **5:** 455-462.

Cohen J.I. 2005. Poorer countries turn to publicly developed GM crops. *Nature Biotechnology*. **23:** 27-33.

Colmenares A.J., Aleu J., Duran-Patron R., Collado I.G., Hernandez-Galan R. 2002. The putative role of botrydial and related metabolites in the infection mechanism of *Botrytis cinerea*. *Journal of Chemical Ecology*. **28:** 997–1005.

Compant S., Clément C. and Sessitsch A. 2010. Plant growth-promoting bacteria in the rhizo- and endosphere of plants: their role, colonisation, mechanisms involved and prospects for utilization. *Soil Biology and Biochemistry*. **42:** 669-678.

Compant S., Mitter B., Colli-Mull J.G., Gangl H. and Sessitsch A. 2011. Endophytes of grapevine flowers, berries, and seeds: identification of cultivable bacteria, comparison with other plant parts, and visualization of niches colonization. *Microbial Ecology*. **62:** 188-197.

Conn V. M., and Franco C. M. M. 2004. Analysis of the endophytic actinobacterial population in the roots of wheat (*Triticum aestivum* L.) by terminal restriction fragment length polymorphism and sequencing of 16S rRNA clones. *Applied Environmental Microbiology*. **70:** 1787-1794.

Corbisier P., Broothaerts W., Gioria S., Schimmel H., Burns M., Baoutina A., Emslie K.R., Furui S., Kurosawa Y., Holden M.J., *et al.* 2007. Towards metrological traceability for DNA fragment ratios in GM quantification. 1. Effects of DNA extraction methods on the quantitative determination of Bt176 corn by real-time PCR. *Journal of Agricultural and Food Chemistry*. **55:** 3249-3257.

Crist T.O. and Veech J.A. 2006. Additives partitioning of rarefaction curves and species-area relationships: unifying α-, β- and γ-diversity with sample size and habiata area. *Ecology Letters*. **9:** 923-932.

Culman S.W., Bukowski R., Gauch H.G., Cadillo-Quiroz H. and Buckley D.H. 2009. T-REX: software for the processing and analysis of T-RFLP data. *BMC Bioinforma*. **10:** 177.

Cutler D.F., Botha T. and Stevenson D.Wm. 2007. Plant anatomy: an applied approach. Blackwell Publishing. Australia. 1st Edition. Chapter 2-6. Pages 14-108.

De Lipthay J.R., Enzinger C., Johnsen K. Aamand J. and Sorensen S.J. 2004. Impact of DNA extraction method on bacterial community composition measured by denaturing gradient gel electrophoresis. *Soil Biology and Biochemistry*. **36:** 1607-1614.

Demeke T. and Jenkins G.R. 2010. Influence of DNA extraction methods, PCR inhibitors and quantification methods on real-time PCR assay of biotechnology-derived traits. *Analysis of Bioanalytical Chemistry*. **396:** 1977-1990.

Demeke T. and Ratnayaka I. 2009. Effects of DNA extraction and purification methods on real-time quantitative PCR analysis of roundup soybean. *Journal of AOAC International*. **92:** 1136-1144.

Devi M.C. and Reddy M.N. 2002. Phenolic acid metabolism of groundnut (*Arachis hypogaea* L.) plants inoculated with VAM fungus and *Rhizobium*. *Plant Growth Regulation*. **37:** 151-156.

De Weert S, Vermeiren H., Mulders H.M., Kuipers I., Hendrickx N., Bloemberg G.V., Vanderleyden J., de Mot R. and Lugteberg B.J.J. 2002. Flagella-driven chemotaxis toward exudate components is an important trait for tomato root colonization by *Pseudomonas flourescens*. *Molecular Plant Microbe Interactions*. **15:** 1173-1180.

Ding T., Palmer M.W. and Melcher U. 2013. Community terminal restriction fragment length polymorphisms reveal insights into the diversity and dynamics of leaf endophytic bacteria. *BMC Microbiology.* **13:** 1-11.

Dodd I.C., Zinovkina N.Y., Safronova V.I. and Belimov A.A. 2010. Rhizobacterial mediation of plant hormone status. *Annals of Applied Biology*. **157:** 361-379.

Drabkova L., Kirschner J. and Vlečk Č. 2002. Comparison of seven extraction and amplification protocols in historical herbarium specimens of Juncaceae. *Plant Molecular Biology Reporter*. **20:** 161-175.

Edgar R.C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*. **26:** 2460-2461.

Edgar R.C., Haas B.J.,Clemente J.C., Quince C. and Knight R. 2011. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*. **27:** 2194-2200.

Environmental Protection Agency (EPA), United States. 2004. Carbaryl Ired Facts. Available on: http://www.epa.gov/oppsrrd1/REDs/carbaryl_ired.pdf. Accessed: 09 May 2014.

Enya J., Shinohara H., Yoshida S., Tsukiboshi T., Negishi H., Suyama K. and Tsushima S. 2007. Culturable leaf-associated bacteria on tomato plants and their potential as biological control agents. *Microbial Ecology*. **53:** 524-536.

Epelde L., Mijangos I., Becerril J.M. and Garbisu C. 2009. Soil microbial community as bioindicator of the recovery of soil functioning derived from metal phytoextraction with sorghum. *Soil Biology and Biochemistry*. **41:** 1788-1794.

Farrelly V., Rainey F.A. and Stackebrandt E. 1995. Effect of genome size and rrn gene copy number on PCR amplification of 16S rRNA genes from a mixture of bacterial species. *Applied Environmental Microbiology*. **61:** 2798-2801.

Faure D., Vereecke D., Leveau J.H.J. 2008. Molecular communication in the rhizosphere. *Plant Soil*. **321:** 279-303.

Feng Y.J. Shen D.L. and Song W. 2006. Rice endophyte *Pantoea agglomerans* YS19 promotes host plant growth and affects allocations of host photosynthates. *Journal of Applied Microbiology*. **100:** 938–945.

Fierer N. 2007. Tilting at windmills: a response to a recent critique of terminal restriction fragment polymorphism data. *Applied and Environmental Microbiology*. **73:** 8041-8042.

Fierer N. and Jackson R.B. 2006. The diversity and biogeography of soil bacterial communities. *PNAS*. **103:** 626 631.

Fierer N., Schimel J.P. and Holden P.A. 2003. Influence of drying-rewetting frequency on soil bacterial community structure. *Microbial Ecology*. **45:** 63-71.

Floyd R. Abebe E., Papert A. and Blaxter M. 2002. Molecular barcodes for soil nematode identification. *Molecular Ecology*. **11:** 839–850.

Forchetti G., Masciarelli O., Alemano S., Alvarez D. and Adbala. 2007. Endophytic bacteria in sunflower (*Helianthus annuus* L.) isolation, characterization, and production of jasmonates and abscisic acid in culture medium. *Applied Microbiology and Biotechnology*. **76:** 1145-1152.

Franche C., Lindström K. and Elmerich C. 2009. Nitrogen-fixing bacteria associated with leguminous and non-leguminous plants. *Plant Soil*. **329:** 35-59.

Fredericks D.N., Smith C. and Meier A. 2005. Comparison of six DNA extraction methods for recovery of fungal DNA as assessed by quantitative PCR. *Journal of Clinical Microbiology*. **43:** 5122-5128.

Fürnkranz M. Lukesch B. and Müller H. 2012. Microbial diversity inside pumpkins: microhabitat-specific communities display a high antagonistic potential against phytopathogens. *Microbial Ecology*. **63:** 418-428.

Gagne-Bourgue F., Aliferis K.A., Seguin P., Rani M., Samson R. and Jabaji S. 2013. Isolation and characterisation of indigenous endophytic bacteria associated with the leaves of switchgrass (*Panicum virgatum* L.) cultivars. *Applied Microbiology*. **114:** 836-853.

Gilbert E.J. 1993. *Pseudomonas* lipases: biochemical properties and molecular cloning. *Enzyme and Microbial Technology*. **15:** 634-645.

Gilbert J.A. and Hughes M. 2011. Gene expression profiling: metatranscriptomics. *Methods in Molecular Biology*. **733:** 195-205.

Gillis M. and De Ley J. 2006. The Genera *Chromobacterium* and *Janthinobacterium*. Prokaryotes. **5:** 737-746.

Glazenbrook J. 2005. Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annual Review of Phytopathology*. **43:** 205-227.

Goggin F.L. 2007. Plant-aphid interactions: molecular and ecological perspectives. *Current Opinion in Plant Biology*. **10:** 399-408.

Gomez-Alvarez V., Revetta R.P. and Domingo J.W.S. 2012. Metagenomic analyses of drinking water receiving different disinfection treatments. *Applied and Environmental Microbiology*. **78:** 6095-6102.

Gopalakrishnan S., Humayun P., Kiran B.K., Kannan I.G.K., Vidya M.S., Deepthi K. and Rupela O. 2011. Evaluation of bacteria isolated from rice rhizosphere for biological control of charcoal rot of sorghum caused by *Macrophomina phaseolina* (Tassi) Goid. *World Journal of Microbiology and Biotechnology*. **27:** 1313-1321.

Gotelli N.J. and Colwell R.K. 2011. Estimating species richness. *In*: Magurran A.E. and McGill B.J. (editors). Frontiers in Measuring Biodiversity. New York: Oxford University Press. Pages 39–54.

Gottel N.R., Castro H.F., Kerley M., Yang Z., Pelletier D.A., Podar M., Karpinets T., Uberbacher E., Tuskan G.A., Vilgalys R., *et al.* 2011. Distinct microbial communities within the endosphere and rhizosphere of *Populus deltoides* roots across contrasting soil types. *Applied and Environmental Microbiology*. **77:** 5934-5944.

Gouse M., Pray C.E., Kirsten J. and Schimmelpfenning. 2005. A GM subsistence in Africa: the case of Bt white maize in South Africa. *International Journal of Biotechnology*. **7:** 84-94.

Govan J.R. and Deretic V. 1996. Microbial pathogenesis in cystic fibrosis: mucoid *Pseudomonas aeruginosa* and *Burkholderia cepacia*. *Microbiology and Molecular Biology Reviews*. **60:** 539-574.

Green M.J., Thompson D.A. and MacKenzie D.J. 1999. Easy and efficient DNA extraction protocol from woody plants for the detection of phytoplasmas by polymerase chain reaction. *Plant Disease*. **83:** 484-485.

Griffiths R.I., Whiteley A.S., O'Donell A.G. and Bailey M.J. 2003. Physiological and community responses of established grassland bacterial populations to water stress. *Applied Environmental Microbiology*. **69:** 6961-6968.

Grönemeyer J.L., Burbano C.S., Hurek T. and Reinhold-Hurek B. 2012. Isolation and characterization of root-associated bacteria from agricultural crops in the Kavango region of Namibia. *Plant Soil*. **356:** 67-82.

Gunatilaka A.A.L. 2006. Natural products from plant-associated microorganisms: distribution, structural diversity, bioactivity, and implication of their occurrence. *Journal of Natural Products*. **69:** 509-526.

Guncheva M. and Zhiryakova D. 2011. Catalytic properties and potential applications of *Bacillus* lipases. *Journal of Molecular Catalysis B: Enzymatic*. **68:** 1-21.

Gupta R., Gupta N. and Rathi P. 2004. Bacterial lipases: an overview of production, purification and biochemical properties. *Applied Microbiology and Biotechnology*. **64:** 763-781.

Hallmann J., Quadt-Hallmann A., Mahaffee W.F. and Kloepper J.W. 1997. Bacterial endophytes in agricultural crops. *Canadian Journal of Microbiology*. **43:** 895-914.

Hameeda B., Rupela O.P., Reddy G. and Satyavani K. 2006. Application of plant growth-promoting bacteria associated with compost and macrofauna for growth promotion. *Biology and Fertility of Soils*. **43:** 221-227.

Han J., Choi H., Lee S., Orwin P.M., Kim J., LaRoe, S.L., Kim T., O'Neil J., Leadbetter J.R., Lee S.Y., *et al.* 2011. Complete genome sequencing of the metabolically versatile plant growth-promoting endophyte *Variovorax paradoxus* S110. *Journal of Bacteriology*. **193:** 118-1190.

Hansel C.M., Fendorf S., Jardine P.M. and Francis C.A. 2008. Changes in bacterial and archaeal community structure and functional diveristy along a geochemically variable soil profile. *Applied and Environmental Microbiology*. **74:** 1620-1633.

Hansen M.C., Tolker-Nielsen T., Givskov M. and Molin S. 1998. Biased rDNA PCR amplification caused by interference from DNA flanking the template region. *FEMS Microbiology Ecology*. **26:** 141-149.

Hardoim P.R., Hardoim C.C.P., van Overbeek L.S. and van Elsas J.D. 2012. Dynamics of seed-borne rice endophytes on early growth stages. *PloS ONE*. **7:** e30438.

Hardoim P.R., Van Overbeek L.S. and van Elsas J.D. 2008. Properties of bacterial endophytes and their proposed role in plant growth. *Trends in Microbiology*. **16:** 463-471.

Hartmann M. and Widmer F. 2008. Reliability for detecting composition and changes of microbial communities by T-RFLP genetic profiling. *FEMS Microbiology Ecology*. **63:** 249-260.

Hasan-Beikdashti M., Forootanfar H., Safiarian M.S., Ameri A., Ghahremani M.H., Khoshayand M.R. and Faramarzi. 2012. Optimization of culture conditions for production of lipase by a newly isolated bacterium *Stenotrophomonas maltophilia*. *Journal of the Taiwan Institute of Chemical Engineers*. **43:** 670-677.

He X., Chang W., Pierce D.L., Seib L.O., Wagner J. and Fuqua C. 2003. Quorum sensing in *Rhizobium* sp. strain NGR234 regulates conjugal transfer (*tra*) gene expression and influences growth rate. *Journal of Bacteriology*. **185:** 809-822.

Hellmich R. L. and Hellmich K. A. 2012. Use and Impact of Bt Maize. *Nature Education Knowledge*. **3:** 4.

Heydari A. and Pessarakli M. 2010. A review on biological control of fungal plant pathogens using microbial anatagonists. *Journal of biological Sciences*. **10:** 273-290.

Hirsh A.M. 2004. Plant-microbe symbioses: a continuum from commensalism to parasitism. *Symbiosis*. **37:** 1-19.

Hirsh P.R. Mauchline T.H. and Clark I.M. 2010. Culture-independent molecular techniques for soil microbial ecology. *Soil Biology and Biochemistry*. **42:** 878-887.

Hodge S. and Andrews M. 2004. Factors affecting interspecific facilitation of crop plants by mycorrhizal fungi. *Aspects of Applied Biology*. **72**: 111–124.

Höflich G., Wiehe W. and Kühn G. 1994. Plant growth stimulation by inoculation with symbiotic and associative rhizosphere microorganisms. *Experientia*. **50:** 897-905.

Hoffman M.T. and Arnold A.E. 2010. Diverse bacteria inhabit hyphae of phylogenetically diverse fungal endophytes. *Applied and Environmental Microbiology*. **76:** 4063-4075.

Holm S. 1979. A simple sequentially rejective multiple test procedure. *Scandanavian Journal of Statistics*. **6:** 65-70.

Hongoh Y., Yuzawa H., Ohkuma M. and Kudo T. 2003. Evaluation of primers and PCR conditions for the analysis of 16S rRNA genes from a natural environment. *FEMS Microbiology Letters*. **221:** 299-304.

Hopkins W.G., Norman P. and Hüner P.A. 2004. Introduction to plant physiology. John & Wiley Sons. Inc. USA. 3rd Edition. Chapter 1, Pages 17-27.

Horne E.C., Kumpatla S.P., Patterson K.A., Gupta M. and Thompson S.A. 2004. Improved high-throughput sunflower and cotton genomic DNA extraction and PCR fidelity. *Plant Molecular Biology Reporter*. **22:** 83a-83i.

Hurek T. Egener T. and Reinhold-Hurek B. 1997. Divergence in nitrogenases of *Azoarcus* spp., Proteobacteria of the β subclass. *Journal of Bacteriology*. **179:** 4172-4178.

Hurek T. Handley L., Reinhold-Hurek B. and Piche Y. 1998. Does *Azoarcus* sp. fix nitrogen with monocots? *In* Elmerich, C., Kondorosi, A., Newton, W.E. (editors), Biological Nitrogen Fixation for the 21st Century. Kluwer, Dordrecht. Page 407.

Hurek T., Handley L., Reinhold-Hurek B. and Piche Y. 2002. *Azoarcus* grass endophytes contribute fixed nitrogen to the plant in an unculturable state. *MPMI*. **15:** 233-242.

Hurek T., Reinhold-Hurek B., Van Montagu M. and Kellenberg E. 1994. Root colonization and systemic spreading of *Azoarcus* sp. strain BH72 in grasses. *Journal of Bacteriology*. **176:** 1913-1923.

International Crops Research Institute for the Semi-Arid Tropics (ICRISAT). www.icrisat.org.

Ìnceoğlu Ö., Al-Soud W.A., Salles J.F., Alexander V., Semenov A.V., and van Elsas J.D., 2011. Comparative analysis of bacterial communities in a potato field as determined by pyrosequencing. *PlosOne*. **6:** 1-11.

Jacobs B.F., Kingston J.D. and Jacobs L.L. 1999. The origin of grass-dominated ecosystems. *Annals of the Missouri Botanical Garden*. **86:** 590-643.

Jaeger K. and Eggert T. 2002. Lipases for biotechnology. *Current Opinion in Biotechnology*. **13:** 390-397.

Jaimini D., Shabnam A.A. and Sarkar C. 2012. *In-silico* feasibility of novel biodegradation pathways for 1-naphthyl methylcarbamate. *American-Eurasian Journal of Toxicology Sciences*. **4:** 89-93.

James E.K. 2000. Nitrogen fixation in endophytic and associative symbiosis. *Field Crop Research*. **65:** 197-209.

James E.K. and Olivares F.L. 1997. Infection and colonization of sugar cane and other graminaceous plants by endophytic diazotrophs. *Critical Reviews in Plant Sciences*. **17:** 77-119.

James E.K., Olivares F.L., Baldani J.I., Dobereiner J. 1997. *Herbaspirillum*, an endophytic diazotroph colonizing vascular tissue in leaves of *Sorghum bicolor* L. *Moench.* Journal of Experimental Botany. **48:**785-797.

Janda J.M. and Abbott S.L. 2007. 16S rRNA gene sequencing for bacterial identification in the laboratory: pluses, perils and pitfalls. *Journal of Clinical Microbiology*. **45:** 2761-2764.

Janssen P.H. 2006. Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and rRNA genes. *Applied and Environmental Microbiology*. **72:** 1719-1728.

Jeswani H. and Mukherji S. 2013. Batch studies with *Exiguobacterium aurantiacum* degrading structurally diverse organic compounds and its potential for treatment of biomass gasification wastewater. *International Biodeterioration & Biodegradation*. **80:** 1-9.

Jha P. and Kumar A. 2009. Characterization of novel plant growth promoting endophytic bacterium Achromobacter xylosoxidans from wheat plant. *Microbiology Ecology*. **58:** 179-188.

Joo G., Kim Y., Lee I., Song K. and Rhee I. 2004. Growth promotion of red pepper plug seedlings and the production of gibberellins by *Bacillus cereus*, *Bacillus macroides* and *Bacillus pumilus. Biotechnology Letter*. **26:** 487-491.

Justel A., Peña D. and Zamar R. 1997. A multivariate Kolmogorov-Smirnov test of goodness of fit. *Statistics and Probability Letters*. **35:** 251-259.

Kalyuzhnaya M.G., De Marco P., Bowerman S., Pacheco C.C., Lara J.C., Lidstrom M.E. and Christoserdova L. 2006. *Methyloversatilis universalis* gen. nov., sp. nov., a novel taxon within the Betaproteobacteria represented by three methylotrophic isolates. *International Journal of Systematic and Evolutionary Microbiology*. **56:** 2517-2522.

Kamar R., Gohar M., Jéhanno I., Réjasse A., Kallassy M., Lereclus D., Sanchs V. and Ramarao N. 2013. Pathogenic potential of *Bacillus cereus* strains as revealed by phenotypic analysis. *Journal of Clinical Microbiology*. **51:** 320-323.

Kanawaga T. 2003. Bias and artifacts in multitemplate polymerase chain reactions (PCR). *Journal of Bioscience and Bioengineering*. **96:** 317-323.

Kellogg E.A. 2006. Beyond taxonomy: prospects for understanding morphological diversity in the grasses (Poaceae). *Darwiniana*. **44:** 7-17.

Kennedy A.C. and Smith K.L. 1995. Soil microbial diversity and the sustainability of agricultural soils. *Plant and Soil*. **170:** 75-86.

Khanuja S.P.S., Shasany A.K., Darokar M.P. and Kumar S. 1999. Rapid isolation of DNA from dry and fresh samples of plants producing large amounts of secondary metabolites and essential oils. *Plant Molecular Biology Reporter*. **17:** 1-7.

King M. and Oudolf P. 1998. True Grasses: The Gramineae Family, *In* Gardening With Grasses. Francis Lincoln Limited, London. Pages 15-16.

Kirchhof G. Reis V.M., Baldani J.I., Eckert B., Dobereiner J. and Hartmann A. 1997. Occurrence, physiological and molecular analysis of endophytic diazotrophic bacteria in gramineous energy plants. *Plant and Soil*. **194:** 45-55.

Kit Y.S. and Chandran S. 2010. A simple, rapid and efficiend method of isolating DNA from Chokanan mango (*Mangifera indica* L.) *African Journal of Biotechnology*. **9:** 5805-5808.

Kirchhof G. Reis V.M., Baldani J.I., Eckert B., Dobereiner J. and Hartmann A. 1997. Occurrence, physiological and molecular analysis of endophytic diazotrophic bacteria in gramineous energy plants. *Plant and Soil*. **194:** 45-55.

Kraiser T., Gras D.E., Gutiérrez A.G., González B. and Gutiérrez R.A. 2011. A holistic view of nitrogen acquisition in plants. *Journal of Experimental Botany*. **62:** 1455-1466.

Krause A., Ramakumar A., Bartels D., Battistoni F., Bekel T., Boch J., Böhm M., Friedrich F., Hurek T., Krause L., *et al.* 2007. Complete genome of the mutualistic, N2-fixing grass endophyte *Azoarcus* sp. BH72. *Nature Biotechnology.* **25:** 478.

Krechel A., Faupel A., Hallmann J., Ulrich A. and Berg G. 2002. Potato-associated bacteria and their antagonistic potential towards plant-pathogenic fungi and the plant-parasitic nematode *Meloidogyne incognita*. *Canadian Journal of Microbiology*. **48:** 772-786.

Krishna K.R. and Philip L. 2009. Biodegradation of mixed pesticides by mixed pesticide enriched cultures. *Journal of Environmental Science and Health Part B*. **44:** 18.30.

Kroupitski Y., Pinto R., Belausov E. and Sela S. 2011. Distribution of *Salmonella typhimurium* in romaine lettuce leaves. *Food Microbiology*. **28:** 990-997.

Kuklinsky-Sobral J., Araújo W.L., Mendes R., Geraldi I.O., Pizzirani-Kleiner A.A. and Azevedo J.L. 2004. Isolation and characterization of soybean-associated bacteria and their potential for plant growth promotion. *Enviromental Microbiology*. **6:** 1244-1251.

Kusari S., Hertweck C. and Spiteller M. 2012. Chemical Ecology of endophytic fungi: origins of secondary metabolites. Chemistry and Biology. **19:** 792-798.

Lahiri D.K. and Schnabel B. 1993. DNA isolation by a rapid method from human blood samples: effects of MgCl$_2$, EDTA, storage, time and temperature on DNA yield and quality. *Biochemical Genetics*. **31:** 321-328.

Lal R., Dadhwal M., Kumari K., Sharma P., Singh A., Kumari H., Jit S., Gupta S.K., Nigam A., Lal D., Verma M., Kaur J., Bala K. and Jindal S. 2008. *Pseudomonas* sp. to *Sphingobium indicum*: a journey of microbial degradation and bioremediation of hexachlorocyclohexane. *Indian Journal of Microbiology*. **48:** 3-18.

Lednicka D., Mergaert J., Cnockaert M.C. and Swings J. 2000. Isolation and identification of cellulolytic bacteria involved in the degradation of natural cellulosic fibres. *Systematic and Applied Microbiology*. **23:** 292-299.

Lichtfouse E., Navarrete M., Debauke P., Souchere V., Alberola C. and Menassieu J. 2009. Agronomy for sustainable agriculture: a review. *Sustainable Agriculture*. **1:** 17.

Loganathan P. and Nair S. 2004. *Swaminathania salitolerans* gen. nov., sp. nov., a salt-tolerant, nitrogen-fixing and phosphate-solubilizing bacterium from wild rice (*Porteresia coarctata* Tateoka). *International Journal of Systematic and Evolutionary Microbiology*. **54:** 1185-1190.

Lucero M.E., Unc A., Cooke P., Dowd S. and Sun S. 2011. Endophytic microbiome diversity in micropropagated *Atriplex canescens* and *Atriplex torreyi* var *griffithsii*. *PlosOne*. **6:** e16793.

Lugtenberg B. and Kamilova F. 2009. Plant-growth-promoting rhizobacteria. *The Annual Review of Microbiology*. **63:** 541-556.

Lugtenberg B.J.J., Dekkers L. and Bloemberg G.V. 2001. Molecular determinants of rhizospheric colinzation by Pseudomonads. *Annual Review of Phytopathology.* **39:** 461-490.

Luna M.F., Galar M.L., Aprea J., Molinari M.L. and Boiardi J.L. 2010. Colonization of sorghum and wheat by seed inoculation with *Gluconacetobacter diazotrophicus.* *Biotechnology Letters.* **32:** 1071-1076.

Luo S., Xu T., Chen L., Chen J., Rao C., Xiao X., Wan Y., Zeng G., Long F., Liu C. and Liu Y. 2011. Endophyte-assisted promotion of biomass production and metal-uptake of energy crop sweet sorghum by plant-growth-promoting endophyte *Bacillus* sp. SLS18. *Applied Microbiology and Biotechnology.* **93:** 1745-1753.

Lozupone C. and Knight R. 2005. Unifrac: a new phylogenetic method for comparing microbial communities. *Applied Environmental Microbiology.* **71:** 8228-8235.

Lozupone C.A., Hamady M. and Knight R. 2006. UniFrac - An online tool for comparing microbial community Diversity in a Phylogenetic Context. *BMC Bioinformatics.* **7:** 371.

Lozupone C.A., Hamady M., Kelley S.T. and Knight R. 2007. Quantitative and qualitative β diversity measures lead to different insights into factors that structure microbial diversity. *Applied and Environmental Microbiology.* **73:** 1576-1585.

Mace E.S., Buhariwalla H.K. and Crouch J.H. 2003. A high-throughput DNA extraction protocol for tropical molecular breeding programs. *Plant Molecular Biology Reporter.* **21:** 459a-459h.

Madigan M.T., Martinko J.M., Dunlap P.V. and Clark D.P. 2009. Chemistry of cellular components. *In* Brock Biology of Microorganisms. Pearson Education Inc. U.S.A. 12th Edition. Chapter 3, page 57.

Magnani G.S., Didonet C.M., Cruz L.M., Picheth C.F., Pedrosa F.O. and Souza E.M. 2010. Diversity of endophytic bacteria in Brazilian sugarcane. *Genetic and Molecular Research*. **9:** 250-258.

Magnani G.S., Cruz L.M., Weber H., Bespalhok J.C., Daros E., Baura V., Yates M.G., Monteiro R.A., Faoro H., Pedrosa F.O. and Souza E.M. 2013. Culture-independent analysis of endophytic bacterial communities associated with Brazilian sugarcane. *Genetic Molecular Research*. **12:** 4549-4558.

Malfanova N., Kamilova F., Validov S., Shcherbakov A., Chebotar V., Tikhonovich I. and Lugtenberg B. 2011. Characterisation of *Bacillus subtilis* HC8, a novel plant-beneficial endophytic strain from giant hogweed. *Microbial Ecology*. **4:** 523-532.

Mansfield J., Genin S., Magori S., Citovsky V., Sriariyanum M., Ronald P., Dow M., Verdier V., Beer S.V., Machado M.A., *et al.* 2012. Top 10 plant pathogenic bacteria in molecular plant pathology. *Molecular Plant Pathology*. **13:** 614-629.

Marchesi J.R., Sato T., Weightman A.J., Martin T.A., Fry J.C., Hiom S.J. and Wade W.G. 1998. Design and evaluation of useful bacterium-specific PCR primers that amplify genes coding for bacterial 16S rRNA. *Applied and Environmental Microbiology*. **64:** 795-799.

Marques A.P.G.C., Pires C., Moreira H., Rangel A.O.S.S. and Castro P.M.L. 2010. Assessment of the plant growth promotion abilities of six bacterial isolates using *Zea mays* as indicator plant. *Soil Biology & Biochemistry*. **42:** 1229-1235.

Marschner P., Baumann K. and Solaiman Z. 2005. Molecular approaches to study the microbial community structure and function in the rhizosphere. *In* Varma A. and

Podila G.P. (editors). Biotechnological Applications of Microorganisms. I.K. International (Pvt). Ltd. New Delhi. 1st Edition. Pages: 311-334.

Marulanda A., Barea J.M. and Azcón R. 2006. An indigenous drought-tolerant strain of *Glomus intraradices* associated with a native bacterium improves water transport and root development in *Retama sphaerocarpa*. *Microbial Ecology*. **52:** 670-678.

Marulanda A., Barea J.M. and Azcón R. 2009. Stimulation of plant growth and drought tolerance by native microorganisms (AM fungi and bacteria) from dry environments: mechanisms related to bacterial effectiveness. *Journal of Plant Growth Development*. **28:** 115-124.

Matiru V.N. and Dakora F.D. 2004. Potential use of rhizobial bacteria as promoters of plant growth for increased yield in landraces of African cereal crops. *African Journal of Biotechnology*. **3:** 1-7.

Mattila P. and Hellström J. 2007. Phenolic acids in potatoes, vegetables, and some of their products. *Journal of Food Composition and Analysis*. **20:** 152-160.

McDonald D., Price M.N., Goodrich J., Nawrocki E.P., DeSantis T.Z., Probst A., Andersen G.L., Knight R., Hugenholtz P. 2012. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME Journal.* **6:** 610–618.

McGuinness M. and Dowling D. 2009. Plant-associated bacterial degradation of toxic organic compounds in soil. *International Journal of Environmental Research and Public Health*. **6:** 2226-2247.

McNear Jr. and David H. The rhizosphere – roots, soil and everything in between. *Nature Education Knowledge*. **4:** 1.

Mehnaz S. and Lazarovits G. 2006. Inoculation effects of *Pseudomonas putida*, *Gluconacetobacter azotocaptans*, and *Azospirillum lipoferum* on corn plant growth under greenhouse conditions. *Microbial Ecology.* **3:** 326-335.

Mendes R., Pizzirani-Kleiner A., Araujo W. and Raaijmarkers J. 2007. Diversity of cultivated endophytic bacteria from sugarcane: genetic and biochemical characterization of *Burkholderia cepacia* complex isolates. *Applied and Environmental Microbiology.* **73:** 7259-7267.

Mercado-Blanco J. and Bakker P.A.H.M. 2007. Interactions between plants and beneficial *Pseudomonas* spp.: exploiting bacterial traits for crop protection. *Antonie van Leeuwenhoek.* **92:** 367-389.

Micali C.O., Neumann U., Grunewald D., Panstruga R. and O'Connell R. 2011. Biogenesis of a specialized plant-fungal interface during host cell internalization of *Golovinimyces orontii* haustoria. *Cellular Microbiology.* **13:** 210-226.

Miethke M. and Marahiel M.A. 2007. Siderophore-based iron acquisition and pathogen control. *Microbiology and Molecular Biology Reviews.* **71:** 413-451.

Miguel P.S.B., Delvaux J.C., de Oliveira M.N.V. Monteiro L.C.P., de Souza Freitas F., Costa M.D., Totola M.R., de Moraes C.A. and Borges A.C. 2013. Diversity of endophytic bacteria in the fruits of *Coffea canephora. African Journal of Microbiology Research.* **7:** 586-594.

Miller D.N., Bryant J.E., Madsen E.L. and Ghiorse W.C. 1999. Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples. *Applied and Environmental Microbiology.* **65:** 4715-4724.

Miller S.H., Browne P., Prigent-Combaret C., Combes-Meynet E., Morrissey J.P. and O'Gara F. 2010. Biochemical and genomic comparison of inorganic phosphate solubilisation in *Pseudomonas* species. *Environmental Microbiology Reports*. **2:** 403-411.

Mohammadi K. And Sohrabi Y. 2012. Bacterial biofertilizers for sustainable crop production: a review. *ARPN Journal of Agricultural and Biological Science*. **7:** 307-316.

Mohr T.J., Liu H., Yan S., Morris C.E., Castillo J.A., Jelenska J. and Vinatzer B.A. 2008. Naturally occurring nonpathogenic isolates of the plant pathogen *Pseudomonas syringae* lack a Type III secretion system and effector gene orthologues. *Journal of Bacteriology*. 190: 2858-2870.

Molina L.G., da Fonseca G.C., de Morais G.L., de Oliveira L.F.P., de Carvalho J.B., Kulcheski F.R., and Margis R. 2012. Metatranscriptomic analysis of small RNAs present in soybean deep sequencing libraries. *Genetics and Molecular Biology*. **35:** 292-303.

Montesinos E. 2003. Plant-associated microorganisms: a view from the scope of microbiology. *Internal Microbiology*. **6:** 221-223.

Monteiro J.M., Vollú R.E., Coelho M.R.R., Alviano C.S., Blank A.F. and Seldin L. 2009. Comparison of the bacterial community and characterisation of plant growth-promoting rhizobacteria from different genotypes of *Chrysopogon zizaniodes* (L.) Roberty (Vertiver) rhizospheres. *The Journal of Microbiology*. **47:** 363-370.

Moré M.I., Herrick J.B., Silva M.C., Ghiorse W.C. and Madsen E.L. 1994. Quantitative cell lysis of indigenous microorganisms and rapid extraction of microbial DNA from sediment. *Applied and Environmental Microbiology*. **60**: 1572-1580

Morgan J.A.W., Bending G.D. and White P.J. 2005. Biological costs and benefits to plant-microbe interactions in the rhizosphere. *Journal of Experimental Botany*. **56:** 1729-1739.

Morgan J.A.W and Whipps J.M. 2001. Methodological approaches to the study of rhizosphere carbon flow and microbial population dynamics. *In*: Pinton R, Varanini Z, Nannipieri P. (editors). The rhizosphere: biochemistry and organic substances at the soil–plant interface. New York: Marcel Dekker, pp 373–410.

Mornkham T., Wangsomnuk P.P., Wangsomnuk P., Jogloy S. Pattanothai A. and Fu Y.B. 2012. Comparison of five DNA extraction methods for molecular analysis of Jerusalem artichoke (*Helianthus tuberosus*). *Genetic and Molecular Research*. **11:** 572-581.

Motari M., Quach U., Thorsteinsdottir H., Martin D.K., Daar A.S. and Singer P.A. 2004. South Africa – blazing a trail for African biotechnology. *Nature Biotechnology*. **22:** 37-41.

Murray M. and Thompson W. 1980. Rapid isolation of high molecular weight DNA plant DNA. *Nucleic Acid Research*. **8:** 4321-4325

Muyzer G., De Waal E.C. and Uitierlinden A.G. 1993. Profiling of Complex Microbial Populations by Denaturing Gradient Gel Electrophoresis Analysis of Polymerase Chain Reaction-Amplified Genes Coding for 16S rRNA. *Applied and Environmental Microbiology*. **59:** 695-700.

Mycoroot™, Pty (Ltd). www.mycoroot.com. South Africa.

Natural Industries Inc. www.naturalindustries.com. U.S.A.

Nautiyal C.S. 1998. An efficient microbiological growth medium for screening phosphate solubilizing microorganisms. *FEMS Microbiology Letters* **170**: 265-270.

Neher, D. 1992. Ecological sustainability in agricultural systems. *Journal of Sustainable Agriculture*. **2:** 51-61.

Newton A.C., Fitt B.D.L., Atkins S.D., Walters D.R. and Daniell T.J. 2010. Pathogenesis, parasitism and mutualism in the trophic space of microbe-plant interactions. *Trends in Microbiology*. **18:** 365-373.

Ng P.C. and Kirkness E.F. 2010. Whole genome sequencing. *Methods in Molecular Biology*. **628:** 215-226.

Niu C., Kebede H., Auld D.L., Woodward J.E., Burow G. and Wright R.J. 2008. A safe inexpensive method to isolate high quality plant and fungal DNA in an open laboratory environment. *African Journal of Biotechnology*. **7:** 2818-2822.

Nocker A., Burr M. and Camper A.K. 2007. Genotypic microbial community profiling: a critical technical review. *Microbial Ecology*. **54:** 276-289.

Oh C. and Beer S.V. 2005. Molecular genetics of *Erwinia amylovora* in the development of fire blight. *FEMS Microbiology letters*. **253:** 185-192.

Okubo T., Ikeda S., Yamashita A., Terasawa K. and Minamisawa K. 2012. Pyrosequence read length of 16S rRNA gene affects phylogenetic assignment of plant-associated bacteria. *Microbes and Environments*. **27:** 204-208.

Orole O.O. and Adejumo T.O. 2011. Bacterial and fungal endophytes associated with grains and roots of maize. *Journal of Ecology and the Natural Environment.* **3:** 298-303.

Ottesen A.R., Pena A.G., White J.R., Pettengill J.B., Li C., Allard S., Rideout S., Allard M., Hill T., Evans P., et al. 2013. Baseline survey of the anatomical microbial ecology of an important food plant: *Solanum lycopersicum* (tomato). *BMC Microbiology.* **13:** 114-125.

Park K.H., Lee C.Y. and Son H.J. 2009. Mechanism of insoluble phosphate solubilization by *Pseudomonas fluorescens* RAF15 isolated from ginseng rhizosphere and its plant growth-promoting activities. *Letters in Applied Microbiology.* **49:** 222-228.

Passioura J. 2007. The drought environment: physical, biological and agricultural perspectives. *Journal of Experimental Botany.* **58:** 113-117.

Pereira P., Ibáñez F., Rosenblueth M., Etcheverry M. and Martínez-Romero. 2011. Analysis of the bacterial diversity of the bacterial diversity associated with the roots of maize (*Zea mays* L.) through culture-dependent and culture-independent methods. *ISRN Ecology.* Article ID 938546.

Petrosino J.F., Highlander S., Luna R.A., Gibbs R.A. and Versalovic J. 2009. Metagenomic pyrosequencing and microbial identification. *Clinical Chemistry.* **55:** 856-8666.

Piccinin R.C.R., Ebbs S.D., Reichman S.M., Kolev S.D., Woodrow I.E. and Baker A.J.M. 2007. A screen of some native Australian flora and exotic agricultural species

for their potential application in cyanide-induced phytoextraction of gold. *Minerals Engineering*. **20:** 1327-1330.

Piccoli P., Travaglia C., Cohen A., Sosa L., Cornejo P., Masuelli R. and Bottini R. 2011. An endophytic bacterium isolated from roots of the halophyte *Prosopis strombulifera* produces ABA, IAA, gibberellins A_1 and A_3 and jasmonic acid in chemically-defined culture medium. *Plant Growth Regulation*. **64:** 207-210.

Piest R., Honsel D., Twieling G. and Löffert D. 2001. PCR inhibitors in plant DNA preparations. *Qiagen news*. **3:** 7–9.

Pilon-Smit E. 2005. Phytoremediation. *Annual Review of Plant Biology*. **56:** 15-39.

Pisarska K. And Pietr S.J. 2012. Isolation and partial characterization of culturable endophytic *Arthrobacter* spp. from leaves of maize (*Zea mays* L.). *Communications in Agricultural and Applied Biological Sciences*. **77:** 225-233.

Pitzschke A. 2013. *Agrobacterium* infection and plant defense – transformation success hangs by a thread. *Frontiers in Plant Science*. **4:** 10.3389.

Podile A.R. and Kishore G.K. 2006. Plant-growth-promoting rhizobacteria. *In* Gnanamanickam S.S. (editor). Plant-Associated Bacteria. Springer. Pp 195-230.

Podolich O., Laschevskyy V., Ovcharenko L., Kozyrovska N. and Pirttila A.M. 2009. *Methylobacterium* sp. resides in unculturable state in potato tissue in vitro and becomes culturable after induction by *Pseudomonas fluorescens* IMGB163. *Journal of Applied Microbiology*. **106:** 728-737.

Porebski S., Bailey L.G. and Baum B.R. 1997. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenolic compounds. *Plant Molecular Biology Reporter*. **15:** 8-15.

Price M.N., Dehal P.S. and Arkin A.P. 2009. FastTree: Computing Large Minimum-Evolution Trees with Profiles instead of a Distance Matrix. *Molecular Biology and Evolution*. **26:** 1641-1650.

Prieto P., Schilirò E., Maldonado-González, Valderrama R., Barroso-Albarracín J.B., Mercado-Blanco J. 2011. Root hairs play an important role in the endophytic colonization of olive roots by *Pseudomonas* spp. with biocontrol activity. *Microbial Ecology*. **62:** 435-445.

Prosser J.I. 2010. Replicate or lie. *Environmental Microbiology*. **12:** 1806-1810.

Pupo G.M., Karaolis D.K., Lan R. and Reeves P.R. 1997. Evolutionary relationships among pathogenic and nonpathogenic *Escherichia coli* strains inferred from multilocus enzymes electrophoresis and *mdh* sequence studies. *Infection and Immunity*. **65:** 2685-2692.

Qin S., Xing K., Jiang j., Xu L. and Li W. 2011. Biodiversity, bioactive natural products and biotechnological potential of plant-associated endophytic actinobacteria. *Applied Microbiological Biotechnology*. **89:** 457-473.

R Development Core Team. 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.

Raaijmakers J.M., Leeman M., Van Oorschot M.M.P., Van der Sluis I., Schippers B. and Bakker P.A.H.M. 1995. Dose-response relationships in biological control of fusarium wilt of radish by *Pseudomonas* spp. *Phytopathology*. **85:** 1075-1081.

Raaijmakers J.M., Paulitz T.C., Steinberg C., Alabouvette C. and Moenne-Loccoz Y. 2009. The Rhizosphere: a playground for soilborne pathogens and beneficial microorganisms. *Plant Soil*. **321:** 341-361.

Radianingtyas H., Robinson G.K. and Bull A.T. 2003. Characterization of a soil-derived bacterial consortium degrading 4-chloroaniline. *Microbiology*. **149:** 3279-3287.

Raina V., Suar M., Singh A., Prakash O., Dadhwal M., Gupta S.K., Dogra C., Lawlor K., Lal S. and van der Meer J.R. 2008. Enhanced biodegradation of hexachlorohexane (HCH) in contaminated soils via inoculation with *Sphingobium indicum* B90A. *Biodegradation*. **19:** 27-40.

Raj S.N., Shetty N.P. and Shetty H.S. 2004. Seed bio-priming with *Pseudomonas fluorescens* isolates enhances growth of pearl millet plants and induces resistance against downy mildew. *International Journal of Pest Management*. **50:** 41-48.

Rajkumar M., Ae N. and Freitas H. 2009. Endophytic bacteria and their potential to enhance heavy metal phytoextraction. *Chemosphere*. **77:** 156-160.

Ramond J., Tshabuse F., Bodpa C.W., Cowan D.A. and Tuffin M.I. 2013. Evidence of variability in the structure and recruitment of rhizospheric and endophytic bacterial communities associated with arable sweet sorghum (*Sorghum bicolor* (*L*) Moench). *Plant and Soil*. **372:** 265-278.

Ramadevi R., Rao K.V. and Reddy V.D. 2014. *Agrobacterium tumefaciens*-mediated transformation and production of stable transgenic pearl millet (*Pennisetum glaucum* [L.] R. Br.). *In Vitro Cellular and Developmental Biology*. DOI10.1007.

Ramesh R., Joshi A.A. and Ghanekar M.P. 2009. Pseudomonads: major antagonistic endophytic bacteria to suppress bacterial wilt pathogen, *Ralstonia solanacearum* in the eggplant (*Solanum melongena* L.). *World Journal of Microbiology and Biotechnology.* **25:** 47-55.

Rappé M.S. and Giovannoni S.J. 2003. The uncultured microbial majority. Annual Review of Microbiology. **57:** 369-394.

Rasche F., Lewis H., Zachow C., Berg G., van Elsas J.D. and Sessitch A. 2006. Impact of transgenic potatoes expressing antibacterial agents on bacterial endophytes is comparable with the effect of plant genotype, soil type and pathogen infection. *Journal of Applied Ecology.* **43:** 555-566.

Rashid S., Charles T.C. and Glick B.R. 2012. Isolation and characterization of new plant growth promoting bacterial endophytes. *Applied Soil Ecology.* **61:** 217-224.

Reinhold-Hurek B. and Hurek T. 1998a. Interactions of gramineous plants with Azoarcus spp. and other diazotrophs: identification, localization, and perspectives to study their function. *Critical Reviews in Plant Sciences.* **17:** 29-54.

Reinhold-Hurek B. and Hurek T. 1998b. Life in grasses: diazotrophic endophytes. *Trends in Microbiology.* **6:** 139-144.

Reinhold-Hurek B. and Hurek T. 2011. Living inside plants: bacterial endophytes. *Current Opinion in Plant Biology.* **14:** 435-443.

Reinecke S.A. and Reinecke A.J. 2007. The impact of organophosphate pesticides in orchards in the Western Cape, South Africa. *Ecotoxicology and Environmental Safety.* **26:** 244-251.

Reis V. M., Vanderleyden J., and Spaepen S. 2011. N2-fixing endophytes of grasses and cereals. *In* Ecological Aspects of Nitrogen Metabolism in Plants. Pages 231-253.

Revathi K., Haribabu T.E. and Sudha P.N. 2011. Phytoremediation of chromium contaminated soil using sorghum plant. *Phytoremediation Journal of Environmental Sciences*. **2:** 417-428.

Requena N., Perez-Solis E., Azcón-Aguilar C., Jeffries P. and Barea J.M. 2001. Increased establishment of mycorrhizal plants inoculated with native endophytes in semi-arid degraded mediterranean ecosystems. *Applied and Environmental Microbiology.* **67:** 495-498.

Rillig M.C. 2004. Arbuscular mycorrhizae, glomalin, and soil aggregation. *Canadian Journal of Soil Science*. **84**: 355–363.

Rincon-Florez V.A., Carvalhais L.C. and Schenk P.M. 2013. Culture-independent molecular tools for soil rhizosphere microbiology. *Diversity*. **5:** 581-612.

Risacher T., Craze M., Bowden S., Paul W. and Barsby T. 2009. Highly efficient *Agrobacterium*-mediated transformation of wheat via in planta inoculation. *Methods in Molecular Biology*. **478:** 115-124.

Roche 454 Sequencing: www.454.com.

Rodriguez H. and Fraga R. 1999. Phosphate solubilizing bacteria and their role in plant growth promotion. *Biotechnology Advances*. **17:** 319-339.

Ronalghi M. 2001. Pyrosequencing sheds light on DNA sequencing. *Genome Research*. **11:** 3-11.

Rosenblueth M. and Martínez-Romero E. 2006. Bacterial endophyted and their interactions with hosts. *MPMI*. **19:** 827-837.

Rösner T. and van Schalkwyk A., 2000. The environmental impact of gold mine tailings footprints in the Johannesburg region, South Africa. *Bulletin of Engineering Geology and the Environment*. **59:** 137-148.

Rotthauwe J.H., Witzel K.P. and Liesack W. 1997. The ammonia monooxygenase structural gene *amoA* as a functional marker: molecular fine-scale analysis of natural ammonia-oxidising populations. *Applied and Environmental Microbiology*. **63:** 4704-4715.

Ryan R.P., Germaine K., Franks A., Ryan D.J. and Dowling D.N. 2008. Bacterial endophytes: recent developments and applications. *FEMS Microbiology Letters*. **278:** 1-9.

Saha R., Saha N., Donofrio R.S. and Bestervelt L.L. 2012. Microbial siderophores: a mini review. *Journal of Basic Microbiology*. **53:** 303-317.

Sahu S.K., Thanaraj M. and Kathiresan K. 2012. DNA Extraction protocol for plants with high levels of secondary metabolites and polysaccharides without using liquid nitrogen and phenol. *ISRN Molecular Biology*. **12:** Article ID 205049.

Saikishore N., Visarada K.B.R.S., Rao S.V. and Seetharama N. 2011. Progress and prospects for *Agrobacterium*-mediated genetic tranformation in sorghum in comparison to other cereals. *Transgenic Plant Journal*. **5:** 27-34.

Salles J.F., van Veen J.A. and J.D. van Elsas. 2004. Multivariate analysis of Burkholderia Species in Soil: Effect of Crop and Land Use History. *Applied and Environmental Microbiology*. **70:** 4012-4020.

Salonen A., Nikkilä J., Jalanka-Tuovinen J., Immonen O., Rajilic-Stojanovic M., Kekkonen R.A., Palva A. and de Vos W. 2010. Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: effective recovery of bacterial and archeal DNA using mechanical cell lysis. *Journal of Microbiological Methods*. **81:** 127-134.

Santos S.R. and Ochman H. 2004. Identification and phylogenetic sorting of bacterial lineages with universally conserved genes and proteins. *Environmental Microbiology*. **6:** 754-759.

Schäfer H. and Muyzer G. 2001. Denaturing Gradient Gel Electrophoresis in Marine Microbial Ecology. *Methods in Microbiology*. **30:** 425-468.

Schenk P.M., Carvalhais L.C. and Kazan K. 2012. Unraveling plant-microbe interactions: can multi-species transcriptomics help? *Trends in Biotechnology*. **30:** 177-184.
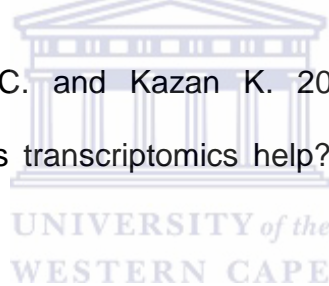
Schloss P.D., Westcott S.L., Ryabin T., Hall J.R., Hartmann M., Hollister E.B., Lesniewski R.A., Oakley B.B. and Parks D.H *et al*. 2009. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*. **75:** 7537-7541.

Schmidt M.A., Balsanelli E., Faoro H., Cruz L.M., Wassem R., de Baura V.A., Weiss V., Yates M.G., Madeira H.M., Pereira-Ferrari L. 2012. The type-III secretion system is necessary for development of a pathogenic and endophytic interaction between *Herbarpirillum rubrisulbalbicans* and Poaceae. *BMC Microbiology*. **12:** 98

Schmieder R. and Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics.* **27:** 863-864.

Schoenholtz S.H., Van Miegroet H. and Burger J.A. 2000. A review of chemical and physical properties as indicators of forest soil quality: challenges and opportunities. *Forest Ecology and Management.* **138:** 335-356.

Scholz W. and Schnabel U. 2006. Decision making under uncertainty in case of soil remediation. *Journal of Environmental Management.* **80:** 132-147.

Schulz R. 2001.Rainfall-induced sediment and pesticide input from orchards into the Lourens River, Western Cape, South Africa: importance of a single event. *Water Research.* **35:** 1869-1876.

Schütte U.M., Abdo Z., Bent S.J., Williams C.J., Pierson J.D. and Forney L.J. 2008. Advances in the use of terminal restriction fragment length polymorphism (T-RFLP) analysis of 16S rRNA genes to characterize microbial communities. *Applied Microbiology and Biotechnology.* **80:** 365-380.

Scupham A.J., Jones J.A., Wesley I.V. 2007. Comparison of DNA extraction methods for analysis of turkey cecal microbiota. *Journal of Applied Microbiology.* **102:** 401–409.

Seghers D., Wittebolle L., Top E.M., Verstraete W. and Siciliano S.D. 2004. Impact of agricultural practices on the *Zea mays* L. endophytic community. *Applied Environmental Microbiology.* **70:** 1475-1482.

Sène M., Gallet C. and Doré T. 2001. Phenolic compounds in a Sahelian sorghum (*Sorghum bicolor*) genotype (CE$_{145\text{-}66}$) and associated soils. *Journal of Chemical Ecology.* **27:** 81-92.

Sessitch A., Gyamfi S., Stralis-Pavese N., Weilharter A. and Pfeifer U. 2002. RNA isolation from soil for bacterial community and functional analysis: evaluation of different extraction and soil conservation protocols. *Journal of Microbiological Methods*. **51:** 171-179.

Sessitsch A., Hardoim P., Döring J., Weilharter A., Krause A., Woyke T., Mitter B., Hauberg-Lotte L., Friedrich F., Rahalkar M., *et al.* 2012. Functional characteristics of an endophyte community colonizing rice roots as revealed by metagenomic analysis. *MPMI*. **25:** 28-36.

Settaluri V.S., Kandala C.V.K., Puppala N. and Sundaram J. 2012. Peanuts and their nutritional aspects-a review. *Food and Nutrition Sciences*. **3:** 1644-1650.

Sevilla M., Burris R.H., Gunapala N., Kennedy C. 2001. Comparison of benefit to sugarcane plant growth and 15N2 incorporation following inoculation of sterile plants with *Acetobacter diazotrophicus* wildtype and *Nif*-mutant strains. *Molecular Plant-Microbe Interactions*. **3:** 358-366.

Sharma P.K., Capalash N., Kaur J. 2007. An improved method for single step purification of metagenomic DNA. *Molecular Biotechnology*. **36:** 61-63.

Sheoran V., Sheoran A.S. and Poonia P. 2013. Phytomining of gold: a review. *Journal of Geochemical Exploration*. **128:** 42-50.

Sheorain V. Banka R. and Chavan M. 2000. Ethanol production from sorghum. *In:* Chandrashekar A., Bandyopadhyay R. and Hall, A.J. (editors).Technical and institutional options for sorghum grain mold management: proceedings of an international consultation, Patancheru (India), 18-19 May 2000. International Crops Research Institute for the Semi-Arid Tropics (ICRISAT). pages 228-239.

Shephard R.N. 1962. The analysis of proximities: multidimensional scaling with an unknown distance function. *Psychometrika*. **27:** 125-140.

Shyu C., Soule T., Bent S.J., Foster J.A. and Forney L.J. 2007. MiCA: a web-based tool for the analysis of microbial communities based on terminal-restriction fragment length polymorphism of 16S and 18S rRNA genes. *Journal of Microbial Ecology*. **53:** 562-570.

Silvani V.A., Fracchia S., Fernandez L., Pergola M. and Godeas A. 2008. A simple method to obtain endophytic microorganisms from field-collected roots. *Soil Biology and Biochemistry*. **40:** 1259-1236.

Skoet J., Stamoulis K. and Deuss A. 2004. Investing in agriculture for growth and food security in the ACP countries. ESA Working Paper No. 04-22. December 2004. Agriculture and Economic Development Analysis Division. The Food and Agriculture Organisation (FAO) of the United Nations.

Smith S.E., Facelli E., Pope S. and Smith F.A. 2010. Plant performance in stressful environments: interpreting new and established knowledge of the roles of arbuscular mycorrhizas*. Plant Soil*. **326:** 3-20.

South African Weather Services(SAWS). www.weathersa.co.za. Accessed 25th September 2012.

Stępień Ł. and Chełkowski J. 2012. Fusarium head blight of wheat: pathogenic species and their mycotoxins. *World Mycotoxin Journal*. **3:** 107-119.

Stewart E.J. 2012. Growing unculturable bacteria. *Journal of Bacteriology*. **194:** 4151-4160.

Stige L.C., Stave J, Chan K., Cianelli L., Pettorelli N. Glantz M. Herren H.R. and Stenseth N.C. 2006. The effect of climate variation on agro-pastoral production in Africa. *PNAS*. **103**: 3049-3053.

Stolz A. 2009. Molecular characteristics of xenobiotic-degrading sphingomonads. *Applied Microbiology and Biotechnology*. **81:** 793-811.

Strobel G. and Daisy B. 2003. Bioprospecting for microbial endophytes and their natural products. *Microbiology and Molecular Biology Reviews*. **67:** 491-502.

Strömberg C.A.E. 2011. Evolution of grasses and grassland ecosystems. *Annals Review of Earth and Planetary Sciences*. **39:** 517-544.

Sun L., Shao H., Zhang R., Zhao L., Li L. and Yao N. 2011. Diversity of siderophore-producing endophytic bacteria of *Cymbidium goeringii* roots. *Acta Microbiologica Sinica*. **51:** 189-195.

Sun L., Qiu F., Zhang X., Dai X., Dong X. and Song W. 2008. Endophytic bacterial diversity in rice (*Oryza sativa* L.) roots estimated by 16S rDNA sequence analysis. *Microbial Ecology*. **55:** 415-424.

Taylor M.P., Van Zyl L., Tuffin M. and Cowan D. 2012. Extremophile and biotechnology: How far have we come? *In* Anitori R.P. (editor). Extremophiles: Microbiology and Biotechnology. Caister Academic Press. Norfolk, UK. Pages: 1-23.

Terrat S., Christen R., Dequiedt S., Leliévre M., Nowak V., Regnier T., Bachar D., Plassart P., Wincker P., Jolivet C. et al. 2012. Molecular biomass and metataxogenomic assessment of soil microbial communities as influenced by soil DNA extraction procedure. *Microbial Biotechnology*. **5:** 135-141.

Terry C.F., Harris N. and Parkes H.C. 2002. Detection of genetically modified crops and their derevatives: critical steps in sample preparation and extraction. *Journal of AOAC International.* **85:** 768-777.

ThermoScientific. Assesement of nucleic acid purity. Technical Bulletin T042. Webpage: www.nanodrop.com/Library. Last viewed: 14 November 2013.

Tien T.M., Gaskins M.H. and Hubbell D.H. 1979. Plant growth substances produced by *Azospirillum brasilense* and their effect on the growth of pearl millet (*Pennisetum americanum* L.). *Applied and Environmental Microbiology*. **37:** 1016-1024.

Tilak K.V.B.R., Ranganayaki N., Pal K.K., De E., Saxena A.K., Nautiyal C.S., Mittal S., Tripathi A.K. and Johri B.N. 2005. Diversity of plant growth and soil health supporting bacteria. *Current Science*. **89:** 135-150.

Tsavkelova E.A., Klimova S.Y., Cherdynsteva T.A. and Netrusov A.I. 2006. Microbial producers of plant growth stimulators and their practical use: a review. *Applied Biochemistry and Microbiology*. **42:** 117-126.

Tunali B., Obanor F., Erginbaş G., Westecott R.A., Nicol J. and Chakraborty S. 2012. Fitness of three *Fusarium* pathogens of wheat. *FEMS Microbiology Ecology*. **81:** 596-609.

Turner S., Pryer K.M., Miao V.P.W. and Palmer, J.D. 1999. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. *Journal of Eukaryotic Microbiology*. **46:** 327–338.

Ulrich K., Ulrich A. and Ewald D. 2008. Diversity of endophytic bacterial communities in poplar grown under field conditions. *FEMS Microbiology Ecology*. **63:** 169-180.

USDA, National Nutrient Database for Standard Reference, Release 25, 2011.

Vamerali T., Bandiera M. and Mosca G. 2010. Field crops for phytoremediation of metal-contaminated land. A review. *Environmental Chemistry Letters.* **8:** 1-17.

Vandenberghe L.P., Rodrigues C., de Olieveira J. and Soccol C.R. 2014. Plant-growth hormones and other phytochemicals. *In*: Brar S.K., Dhillon G.S. and Soccol C.R (editors). Biotransformation of Waste Biomass in High Value Biochemicals. Springer. pages 163-183.

Vanneste S. and Friml J. 2009. Auxin: a trigger for change in plant development. *Cell.* **136:** 1005-1016.

Vargas L., de Carvalho T.L.G., Ferreira P.C.G., Baldani V.L.D. Baldani J.I. and Hemerly A.S. 2012. Early responses of rice (*Oryza sativa* L.) seedlings to inoculation with beneficial diazotrophic bacteria are dependent on plant and bacterial genotypes. *Plant Soil.* **356:** 127-137.
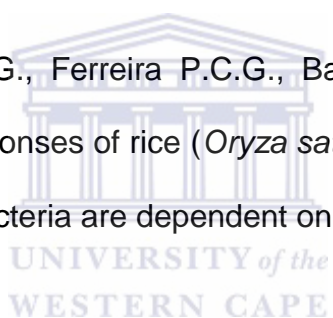
Varma A., Padh H. and Shrivastava N. 2007. Plant genomic DNA isolation: an art or a science. *Biotechology Journal.* **2:** 386-392.

Vartoukian S.R., Palmer R.M. and Wade W.G. 2010. Strategies for culture of "unculturable" bacteria. *FEMS Microbiology Letters.* **309:** 1-7.

Venkateswarlu K., Chendrayan K. and Senthurathan N. 1980. Persistent and biodegradation of carbaxyl in soils. *Journal of Environmental Science and Health*, Part B. **15:** 421-429.

Verma S. C., Singh A., Chowdhury S. P., and Tripathi A. K. 2004. Endophytic colonization ability of two deep-water rice endophytes, *Pantoea* sp. and

Ochrobactrum sp. using green fluorescent protein reporter. *Biotechnology Letters.* **26:** 425-429.

Vessey J.K. 2003. Plant growth promoting rhizobacteria as biofertilizers. *Plant and Soil.* **255:** 571-586.

Vijayalaxmi S. Anu Appaiah K.A., Jayalakshmi S.K., Mulimani V.H. and Sreeramulu K. 2013. Production of bioethanol from fermented sugars of sugarcane bagasse produced by lignocellulolytic enzymes of *Exiguobacterium* sp. VSG-1. *Applied Biochemistry and Biotechnology.* **171:** 246-260.

Wani S.P., Chandrapalaih S., Zambre M.A. and Lee K.K. 1989. Association between N_2-fixing bacteria and pearl millet plants: responses, mechanisms and persistence. *Developments in Plant and Soil Sciences.* **35:** 249-262.

Walker T.S., Bais H.P., Grotewold E. and Vivanco J.M. 2003. Root exudation and rhizosphere biology. *Plant Physiology.* **132:** 44-51.

Watt D.A. and Cramer M.D. 2009. Post-harvest biology of sugarcane. *Sugar Technology.* **11:** 142-145.

Webster J.W. and Akanbi R.T. 2005. Opportunities and challenges for industrial biotechnology in south Africa. Industrial Biotechnology and Biomass Utilisation. United Nations Industrial Development Organisation: Prospects and challenges for the developing world. Stockholm Environment Institute.

Weinstock D.M. and Brown A.E. *Rhodococcus equi*: an emerging pathogen. *Clinical Infectious Diseases.* **34:** 1379-1385.

Werner J.J., Koren O., Hugenholtz P., DeSantis T.Z., Walters W.A., Caporaso J.G., Angenent L.T., Knight R., Ley R.E.. 2012. Impact of training sets on classification of high-throughput bacterial 16S rRNA gene surveys. *ISME Journal.* **6:** 94-103.

West E.R., Cother E.J., Steel C.C. and Ash G.J. 2010. The characterization and diversity of bacterial endophytes of grapevine. *Canadian Journal of Microbiology*. **56:** 209-216.

Weston D.J., Pelletier D.A., Morrell-Falvey J.L., Tschaplinski T.J., Jawdy S.S., Lu T., Allen S.M., Melton S.J., Martin M.Z., Schadt C.W. *et al.*, 2012. *Pseudomonas flourescens* induces strain-dependent and strain-independent host plant responses in defense networks, primary metabolism, photosynthesis and fitness. *MPMI*. **25:** 765-778.

Whipps J.M., Hand P., Pink D. and Bending G.D. 2008. Phyllosphere microbiology with special reference to diversity and plant genotype. *Journal of Applied Microbiology*. **105:** 1744-1755.

Wicoxon F. 1945. Individual comparisons by ranking methods. *Biometrix Bulletin*. **1:** 80-83.

Wieland G., Neumann R. and Backhaus H. 2001. Variation of microbial communities in soil, rhizosphere, and rhizosphere in response to crop species, soil type, and crop development. *Applied and Environmental Microbiology*. **67:** 5849-5854.

Wilfinger W.W., Mackey K., and Chomczynski P.1997. Effect of pH and ionic strength on the spectrophotometric assessment of nucleic acid purity. *BioTechniques.* **22:** 474-481.

Willner D., Daly J., Whiley D., Grimwood K., Wainwright C.E. and Hugenholtz P. 2012. Comparison of DNA Extraction methods for microbial community profiling with an application to pediatric bronchoalveolar lavage samples. *PlosONE*. **7:** e34605.

Wilson I. G. 1997. Inhibition and facilitation of nucleic acid amplification. *Applied and Environmental Microbiology*. **63:** 3741-3751.

Wilson-Corral V., Rodriguez-Lopez M., Lopez-Perez J., Arenas-Vargas M. and Anderson C. 2010. Gold phytomining in arid and semiarid soils. 19th Wolrd Congress of Soil Science, Soil Solutions for a Changing World. 1-6 August 2010, Brisbane, Australia.

Wu D.Y., Ugozzoli L., Pal B.K., Qian J., and Wallace R.B. 1991. The effect of temperature and oligonucleotide primer length on the specificity and efficiency of amplification by the polymerase chain reaction. *DNA and Cell Biology*. **10:** 233-238.

Wu E., Lenderts B., Glassman K., Berezowska-Kaniewska M., Christensen H., Asmus T., Zhen S., Chu U., Cho M. and Zhao Z. 2014.Optimized *Agrobacterium*-mediated sorghum transformation protocol and molecular data of transgenic sorghum plants. *In Vitro Cellular and Developmental Biology*. **50:** 9-18.

Wu T., Chellemi D.O., Graham J.H., Martin K.J. and Rosskopf E.N. 2007. Comparison of soil bacterial communities under diverse agricultural land management and crop production practices. *Microbial Ecology*. **55:** 293-310.

Wuana R.A., Adie P.A., Abah J. and Ejeh M.A. 2013. Screening of pearl millet for phytoextraction potential in soil contaminated with cadmium and lead. *International Journal of Science and Technology.* **2:** 310-319.

Ye J., Singh A. and Ward O.P. 2004. Biodegradation of nitroaromatics and other nitrogen-containing xenobiotics. *World Journal of Microbiology & Biotechnology*. **20:** 117-135.

Yuan S., Cohen D.B., Ravel J., Abdo Z. and Forney L.J. 2012. Evaluation of Methods for the Extraction and Purification of DNA from the Human Microbiome. *PLoS ONE* **7:** e33865.

Zheng X., Liu B., Sun D., Zhu T., Duan P., Xia Y., Ruan C. and Xiao R. 2012. Plant endophyte PLFAs polymorphism in different spatial of citrus HLB. *Chinese Journal of Eco-Agriculture.* **20:** 1-13.

Zehr J.P., Jenkins B.D., Short S.M. and Steward G.F. 2003. Nitrogenase gene diversity and microbial community structure: a cross-system comparison. *Environmental Microbiology*. **5:** 539-554.

Zhou J. Bruns M.A. and Tiedje J.M. 1996. DNA Recovery from soil of diverse composition. *Applied and Environmental Microbiology*. **62**: 316-322.

Zinniel D.K., Lambrecht P., Harris N.B., Feng Z., Kuczmarski D., Higley P., Ishimaru C.A., Arunakumari A., Barletta R.G. and Vidaver A.K. 2002. Isolation and characterisation of endophytic colonizing bacteria from agronomic crops and prairie plants. *Applied and Environmental Microbiology*. **68:** 2198-2208.

Zoetendal E.G., Akkermans A.D.L. and Vos. W.M. 2001. Molecular characterization of microbial communities in the GI-tract based on 16S rRNA sequence diversity. *In*: Dijkhoorn L., Towner K.J. and Struelens M. (editors). New Approaches For The Generation and Analysis of Microbial Typing Data *Science,* Amsterdam. 1st Edition. Chapter 11, pages 267-272.