Characterization of nitrogen-fixing bacteria from *Phaseolus vulgaris* L. in Kenya

by

George Mutegi Mwenda

This thesis is presented for the degree of

Doctor of Philosophy

of

Murdoch University

2017

Declaration

I declare that this thesis is my own account of my research and contains as its main content work which has not previously been submitted for a degree at any tertiary education institution.

....

(George Mutegi Mwenda)

Dedication

I dedicate this 'book' to my daughters, Alexia and Amy, for their immense support during my study.

Abstract

Phaseolus vulgaris (common bean) is an important food crop in Sub-Saharan Africa. Low soil nitrogen limits the productivity of *P. vulgaris* in Kenya and a greater exploitation of symbiotic nitrogen fixation, resulting from interactions with rhizobia, has the potential to improve yields. To enable the increased use of the symbiosis in Kenyan agriculture in the future, studies in this thesis examined the genetic diversity of rhizobia that nodulate *P. vulgaris* in the central and western parts of Kenya, their nitrogen-fixing capabilities, and their competitiveness against *Rhizobium tropici* CIAT 899, a leading commercial inoculant strain for *P. vulgaris*. Lastly, studies investigated the relative importance of the genotypes of resident soil rhizobia, soil rhizobial population densities, inocula densities, and levels of soil nitrogen, in determining nodule occupancy by *R. tropici* CIAT 899 inoculated onto *P. vulgaris*.

Phylogenetic studies using 16S rRNA and *recA* genes indicated at least five species of *Rhizobium* viz., *R. sophoriradicis*, *R. phaseoli*, *R. leucaenae*, *R. paranaense* and *R. etli* nodulate *P. vulgaris* in Kenya. In addition to the five species, strains that likely belong to new species in the genus *Rhizobium* also widely nodulate *P. vulgaris* in Kenyan soils. In glasshouse studies, recovered strains were variably effective on Kenyan cultivars of *P. vulgaris* and 11 fixed as much nitrogen as *R. tropici* CIAT 899. From the 11, strains such as NAK 227, NAK 288, NAK 214 and NAK 157 were also highly competitive in liquid co-inoculation assays, carried out with the aid of *gusA* and *celB* marker genes, and are potential future inoculants for *P. vulgaris* in Kenya. The genotype of the rhizobia in the soil was found to be the primary determinant of the nodule occupancy achieved by the inoculant strain, a finding that conflicts previous reports that indicated nodule occupancy was mainly determined by soil rhizobial densities. The rhizobial genotypes varied in their rhizosphere competence and in their ability to preferentially nodulate the host, suggesting these two traits are important for the successful colonization of *P. vulgaris* nodules by rhizobia.

It is anticipated that future studies will leverage on the results in this thesis, to develop locally-targeted inoculation solutions that optimize nitrogen fixation in *P. vulgaris* in Kenya, and to elucidate the molecular basis for preferential nodulation in *P. vulgaris*.

Table of contents

Declaration .... **i**

Dedication .... **ii**

Abstract .... **iii**

Table of contents .... **iv**

Acknowledgement .... **ix**

Abbreviations .... **x**

CHAPTER 1 Introduction and literature review **1**

1.1 Nitrogen fixation 2

1.2 Legume-rhizobia symbioses 3

1.2.1 Bacterial infection and nodulation 3

1.2.2 Mechanism of N_2 fixation 5

1.2.3 Effectiveness of N_2 fixation 6

1.2.3.1 Causes of variability in N_2 fixation 7

1.2.4 Symbiotic promiscuity 11

1.3 Methods to study taxonomy and diversity of rhizobia 12

1.4 *Phaseolus vulgaris* 15

1.4.1 Cultivation of *P. vulgaris* 15

1.4.2 Effectiveness of *P. vulgaris* symbiosis 16

1.4.3 Diversity of *P. vulgaris*-rhizobia 17

1.4.4 Spread of *P. vulgaris* rhizobia across the globe 18

1.4.5 Rhizobia competition in *P. vulgaris* symbiosis 21

1.4.5.1 Determinants of rhizobial competition outcomes 21

1.4.5.2 Strain typing in competition studies 25

1.5 *P. vulgaris* in Kenya 30

1.6 Aims of thesis 32

CHAPTER 2 The genetic diversity of bacteria that nodulate *P. vulgaris* in Kenya **33**

2.1 Introduction 34

2.2 Materials and methods 35

2.2.1 Nodule collection, strain isolation and storage 35

2.2.2 Strain authentication 37
2.2.3 Genetic characterization 38
2.2.3.1 Genomic DNA preparation 38
2.2.3.2 Genomic DNA fingerprinting by PCR with *nif*-directed RP01 primer 38
2.2.3.3 Restriction Fragment Length Polymorphisms (RFLP)-PCR of 16S rRNA genes 39
2.2.3.4 Amplification and sequencing of 16S rRNA gene 39
2.2.3.5 Amplification and sequencing of *recA* gene 40
2.2.3.6 Amplification and sequencing of *nodC* gene 40
2.3 Results 41
2.3.1 Authenticated strains 41
2.3.2 Diversity of strains based on RP01-PCR 46
2.3.3 Diversity of strains based on RFLP of PCR-amplified 16S rRNA 46
2.3.4 Phylogeny of strains based on 16S rRNA gene 52
2.3.5 Phylogeny of strains based on *recA* gene 54
2.3.6 Phylogeny based on *nodC* genes 56
2.4 Discussion 58
2.4.1 Diverse rhizobia nodulate *P. vulgaris* in Kenya 58
2.4.2 Rhizobia symbiovars nodulating *P. vulgaris* in Kenya 61
2.4.3 Concluding remarks 62
CHAPTER 3 Effectiveness of strains on *P. vulgaris* under controlled conditions 64
3.1 Introduction 65
3.2 Materials and Methods 67
3.2.1 Bacterial strains and *P. vulgaris* cultivar 67
3.2.2 Assessment of N_2 fixation 67
3.2.3 Light and electron microscopy of nodules from variably effective strains 68
3.2.4 Comparing the growth of inoculated *P. vulgaris* cv. KK08 to that supplied with mineral N 69
3.2.5 Data analysis 69
3.3 Results 70
3.3.1 Biomass and N accumulation in shoots of inoculated plants 70
3.3.2 Nodulation in *P. vulgaris* 74
3.3.3 Anatomy of nodules formed by variably effective strains 76
3.3.4 Poly-hydroxybutyrate (PHB) accumulation in bacteroids 77

3.3.5 Link between *nodC* allele and effectiveness 79
3.3.6 Growth of *P. vulgaris* under mineral N in comparison to BNF 80
3.4 Discussion 83
3.4.1 The effectiveness of strains 83
3.4.2 Anatomical features of nodules 84
3.4.3 PHB accumulation 85
3.4.4 *nodC* alleles 85
3.4.5 Ability of BNF to meet N demand in *P. vulgaris* 87
3.4.6 Scoring of nodulation and atypical nodules 88
3.4.7 Concluding remarks 89
CHAPTER 4 Using a broad-host-range plasmid containing *celB* and a *gusA* mini-transposon to assess rhizobial occupancy in *P. vulgaris* nodules 90
4.1 Introduction 91
4.2.1 Bacterial strains, plasmids and media 93
4.2.2 Marking of CIAT 899 with *gusA* 93
4.2.2.1 Transfer of mini-transposons into CIAT 899 93
4.2.2.2 Characterization of marked strains by morphology and growth rate 94
4.2.2.3 N_2 fixation and competitiveness of mutants 95
4.2.2.4 Nodule staining procedure 96
4.2.2.5 Determination of mTn*5*SS*gusA*31 insertion-site in CIAT 899-GE 96
4.2.3 Marking of rhizobia with *celB* gene 99
4.2.3.1 Construction of a stable, broad-host-range CelB plasmid 99
4.2.3.2 Transformation of pGM01 into *E. coli* ST18 100
4.2.3.3 Conjugal transfer of pGM01 into rhizobia 101
4.2.3.4 Confirmation of presence of pGM01 and expression of thermostable CelB 101
4.2.3.5 Characterization of transconjugants by morphology and growth rate 102
4.2.3.6 Plasmid stability in the absence of selection 102
4.2.4 Competitiveness of *celB*-marked strains against CIAT 899-GE for nodulation of *P. vulgaris* 103
4.2.5 Data analysis 104
4.3 Results 105
4.3.1 Marking of *R. tropici* CIAT 899 with *gusA* 105
4.3.1.1. Effect of mini-transposons on growth and morphology of marked strains 105

4.3.1.2 Effect of mini-transposons on N_2 fixation and competitiveness of marked strains 107

4.3.1.3 Insertion site of mTn*5*SS*gusA*31 in the genome of CIAT 899-GE 109

4.3.2 Stable broad-host-range CelB plasmid (pGM01) expressed in rhizobia 113

4.3.2.1 Construction of pGM01 113

4.3.2.2 Conjugation of pGM01 into rhizobia and expression of thermostable β-glucosidase 114

4.3.2.3 Effect of pGM01 on growth rate and morphology of rhizobia 116

4.3.2.4 Stability of pGM01 in rhizobia in the absence of antibiotic selection 117

4.3.3 Competitiveness of strains against CIAT 899-GE for nodulation of *P. vulgaris* 118

4.4 Discussion 122

4.4.1 Dual marker system 122

4.4.2 Competitiveness of strains against CIAT 899 124

4.4.3 Concluding remarks 124

CHAPTER 5 The determinants of inoculation success in *P. vulgaris* 126

5.1 Introduction 127

5.2 Materials and Methods 129

5.2.1 Bacterial strains and *P. vulgaris* cultivar 129

5.2.2 Experiment 1: Rhizosphere competence of strains 130

5.2.3 Experiment 2: Evaluating the effect of four factors on nodule occupancy by seed-applied CIAT 899-GE 131

5.2.4 Data analysis 134

5.3 Results 135

5.3.1 Experiment 1: Rhizosphere competence of strains 135

5.3.2 Experiment 2: Competition between seed-inoculated CIAT 899-GE and soil-borne rhizobia 138

5.3.2.1 Rhizobial population densities in inoculated soil 138

5.3.2.2 Nodule occupancies by seed-inoculated CIAT 899-GE 140

5.3.2.3 Ranking of factors 143

5.4 Discussion 144

5.4.1 Soil rhizobial genotype as the main determinant of nodule occupancy by the inoculant 144

5.4.2 Strain influence on nodule occupancy is effected through rhizosphere competence and preferential nodulation 146

5.4.3 Concluding remarks 148

CHAPTER 6 General discussion **149**
6.0 Recap of aims 150
6.1 Diversity of rhizobia from Kenyan soils 150
6.2 Candidate inoculant strains for Kenya 154
6.3 Limitation of N_2 fixation in *P. vulgaris* 154
6.4 Broad-host-range plasmid for rhizobial competition assays 155
6.5 Determinants of nodule occupancy by the inoculant in *P. vulgaris* 156
Appendices **159**
References **164**

Acknowledgement

The work presented in this thesis has only been possible through the contribution of many. Firstly, I would like to express my sincere gratitude to the N2Africa Project, "Putting nitrogen fixation to work for smallholder farmers in Africa" funded by Bill and Melinda Gates Foundation, and Murdoch University for a PhD scholarship. Many good people at N2Africa contributed to my work in many ways, but special thanks to Prof. Ken Giller for the continuous interest in my work, encouragement and valuable discussions.

I also thank my supervisors, Dr Graham O'Hara, Dr Jason Terpolilli, and Prof. John Howieson, from whom I have garnered a wealth of knowledge. Your enthusiasm, insight, criticism and encouragement were invaluable throughout my PhD journey. Jason, I cannot thank you enough for your guidance, ideas, and for always having 'five minutes' for me even at your busiest times.

I would also like to thank current and former students and staff at the Centre for Rhizobium Studies (CRS), Murdoch, for many informative discussions and presentations over the years. The brilliant people I was privileged to call my colleagues include Aliyu, Amanuel, Chris, Felipe, Jessica, Julie, Liza, Mazvita, Mike, Rachel, Ravi, Rebecca, Regina, Rob, Sofie, Talitha, Tian, Tim, Tom, Yvette, Wani and Wayne. Special thanks to Amanuel Asrat for friendship, advice and good humour and to Dr Sofie De Meyer and Dr Wayne Reeve for their technical inputs into Chapter 2 and 4, respectively.

Many people outside of the CRS helped me along the way. Special thanks to Dr Freddy Baijukya for providing *P. vulgaris* seed, and Prof. Nancy Karanja and Dr Paul Woomer for input during rhizobia collection. My thanks also to Gordon for help with nodule sections and Mellen for help with, among other things, Ms Excel (data validation lists, pivot tables, VLOOKUP, etc.).

Lastly, I would like to acknowledge the support of my family during my study. Alexia and Amy supported me emotionally mostly via FaceTime, my dad Murango Mwenda and mum Priscilla Karigu never lacked an inspirational piece of advice, and my siblings were always there for me (in a quiet, Mwenda way). I owe you all big time.

Abbreviations

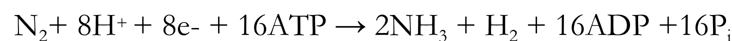
ALA	5-aminolevulinic acid
BLAST	Basic local alignment tool
BNF	Biological nitrogen fixation
bp	Base pair
celB	β-glucosidase gene from *Pyrococcus furiosus*
cv.	Cultivated variety or cultivar
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
FAOSTAT	The Food and Agricultural Organization corporate statistical database
GFP	Green fluorescent protein
GUS	β-glucuronidase
gusA	β-glucuronidase gene
LB	Lysogeny broth
LSD	Least significant difference
MGT	Mean generation time
MLSA	Multilocus sequence analysis
mTn*5*	mini-transposon 5
N	Nitrogen
N_2	Nitrogen gas
OD	Optical density
PCR	Polymerase chain reaction
PHB	Poly-β-hydroxybutyrate
recA	Recombinase A gene
RFLP	Restriction fragment length polymorphism
rpm	Revolutions per minute
rRNA	Ribosomal ribonucleic acid
SDW	Shoot dry weight
sv.	Symbiotic variety or symbiovar
TY	Tryptone yeast
X-Glc	5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

CHAPTER 1
Introduction and literature review

1.1 Nitrogen fixation

Nitrogen is an integral constituent of amino acids, nucleic acids, and many other biologically important organic and inorganic compounds and, therefore, a major factor for the growth of all living things. Although nitrogen (N_2) is abundant in the atmosphere (78% of dry air), the strong triple bond between the nitrogen atoms makes N_2 biologically inaccessible to most organisms. Consequently, for N_2 to be usable for growth, it needs to be converted (fixed) into biologically active forms such as NH_3, NH_4^+ and NO_3^- through artificial or natural processes (Galloway et al., 2004). Artificial N_2 fixation is mainly achieved through the Haber-Bosch process, in which N_2 and H_2 are combined under high pressure, in the presence of a catalyst, to form NH_3 (Smil, 2001). Natural N_2 fixation occurs through lightning and, biologically, through a limited number of N_2–fixing bacteria (Boody & DeVore, 2006; Galloway et al., 2004).

N_2-fixing bacteria are free-living, endophytic, plant-associated, or symbiotic (Herridge et al., 2008; Vitousek et al., 2013). An important group of the symbiotic N_2-fixers is rhizobia, soil bacteria that infect roots of legumes leading to the formation of nodules. Inside the nodules, the bacteria (now called bacteroids) receive carbon and shelter, and in exchange reduce N_2 for the plant (Herridge et al., 2008). The reduction of N_2 into NH_3 by rhizobial bacteroids is achieved through nitrogenase enzyme via the following equation:



$$N_2 + 8H^+ + 8e^- + 16ATP \rightarrow 2NH_3 + H_2 + 16ADP + 16P_i$$

(Dixon & Kahn, 2004)

The reduction of each mole of N_2 is energetically demanding, requiring eight moles of protons, eight moles of electrons, and 16 moles of ATP.

Symbiotic N_2 fixation is a major source of biologically active forms of N for agriculture (Herridge et al., 2008), with approximately 26-27 million tonnes (Tg) of N fixed in 2016 by crop legumes into agricultural systems. This figure is based on projections using an annual increase of 2.44% in the global area under these crops since 2005 (FAOSTAT), when 21 Tg of N was estimated to have been fixed (Herridge et al., 2008). Pasture and fodder legumes fix additional N, estimated to be half of that by crop legumes (Herridge et al., 2008).

At the farm level, symbiotic N_2 fixation has several benefits. Firstly, legumes fix an average of 115 kg N/ha/yr (Herridge et al., 2008) and the additional N substantially increases the yield of legumes and, in some instances, the yields of subsequent crops, through residual N depending on how much of the fixed N is harvested and removed from the agricultural systems (Giller, 2001; Peoples et al., 2009). Secondly, symbiotic N_2 fixation reduces the reliance on fertiliser-N for crop production. Fertiliser-N is costly, and its low efficiency can lead to environmental degradation through gaseous losses, leaching and surface runoffs (Crews & Peoples, 2004; Tilman et al., 2001; Vitousek et al., 1997).

1.2 Legume-rhizobia symbioses

Symbiotic N_2 fixation involving legumes and rhizobia (hereafter simply referred to as N_2 fixation) progresses through a series of chemically controlled events, leading to the development of a root nodule, an organ that provides the micro-aerobic environment conducive for N_2 fixation by rhizobia (Figure 1.1). The leguminous host supplies carbon to the bacteria and in return receives fixed N.

1.2.1 Bacterial infection and nodulation

Legume-rhizobia interactions are initiated by the release of flavonoids and other signalling compounds into the rhizosphere. These molecules trigger the expression of nodulation (*nod*) genes in rhizobia, through transcriptional activators, leading to the synthesis of Nod factors (Oldroyd, 2013). Nod factors are lipo-chitooligosaccharides molecules with an oligomeric backbone of β-1, 4-linked *N*-acetylglucosamine residues carrying modifications at both the reducing and non-reducing ends (Long, 1996; Oldroyd & Downie, 2008). The Nod factors then interact with plant-cell membrane kinases and trigger root hair deformation and calcium oscillations in the nuclei of cortical cells (Limpens et al., 2003; Madsen et al., 2003). The calcium oscillations are consequently perceived by a calcium and calmodulin-dependent protein kinase (CCaMK) which then promotes nodule organogenesis that occurs mainly by mitotic division of cells of the root cortex, resulting in the formation of a nodule primordium (Downie, 2014; Oldroyd et al., 2011).

Concomitant with nodule morphogenesis is the root infection process. Rhizobia infection of legumes occurs through the epidermis, cracks, or root hairs (Boogerd & van

Rossum, 1997; Sprent, 2007). In legumes that undergo root hair infection such as *Phaseolus vulgaris* (common bean) (Tate et al., 1994), Nod factors trigger root hair curling, encapsulating bacteria attached to the root hair surface with the aid of plant lectins and bacterial surface polysaccharides (Oldroyd & Downie, 2008). An invagination of the root hair tip then occurs and bacteria grow down an intracellular tunnel (infection thread), in two or three braided columns that increase in length and push at the tip of the infection thread (Gage, 2002). Most root hair infections occur from the entrapment of cells of one bacterial type by the root hair curl, but occasionally, more than one type may be trapped, leading to double occupied nodules (Gage, 2002). Infection threads are plant derived and can cross cell boundaries, allowing bacteria within to invade cells in the root cortex, and by endocytosis, a plant-derived membrane encapsulates the bacteria within the infected plant cell. It is within these structures called symbiosomes that bacteria differentiate into bacteroids, their N_2 fixing forms (Oldroyd & Downie, 2008). In some legumes, bacteroids transform irreversibly into greatly-enlarged polyploid forms, while in others such as *P. vulgaris,* endoreduplication is absent (Maróti & Kondorosi, 2014). Symbiosomes may contain single or multiple bacteroids (Cermola et al., 2000).

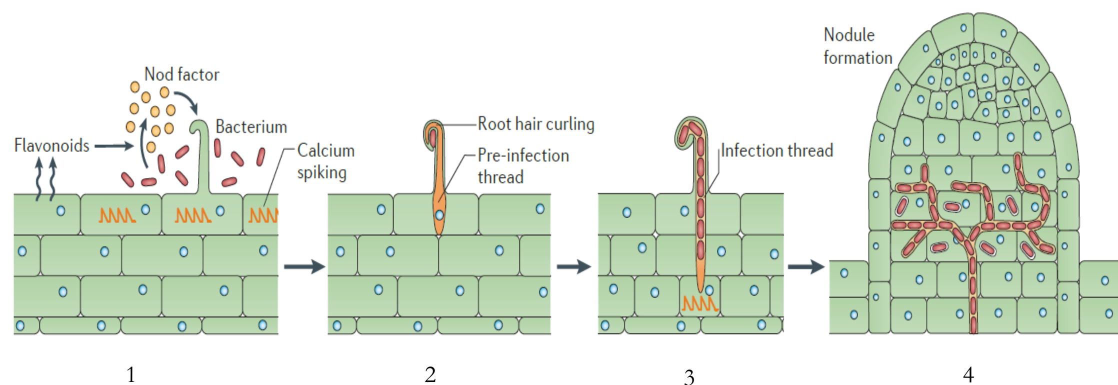


Figure 1.1: A schematic representation of the basic steps involved in the legume nodulation process. In step (**1**) flavonoids excreted by the legume host induce synthesis of Nod factors in rhizobia. Root hairs perceive the Nod factors and trigger calcium spiking and cortical cell divisions that lead to nodule organogenesis. (**2**) Root hair curling encapsulates attached bacteria. (**3**) An invagination of the root tip forms a growing tube filled with dividing bacteria in what is called an infection thread. (**4**) The infection thread branches out and bacteria encapsulated in a symbiosome invade plant cells, differentiate and fix N_2. Image (without modification) from Oldroyd (2013).

Legumes form either indeterminate or determinate nodules, with the type of nodule formed determined by the legume (Ferguson et al., 2010). Indeterminate nodules

originate from the inner cortical cells, and the apical meristem continuously produces new cells that get infected with bacteria, leading to a developmental gradient and cylindrical nodules (Figure 1.2B). Legumes that form indeterminate nodules include *Pisum sativum* (pea), *Medicago* spp., and *Trifolium* spp.

Determinate nodules mostly arise from the outer cortex and have a transient meristem in which synchronous division of infected cells leads to spherical nodules (Figure 1.2A) that last for a few weeks (Cermola et al., 2000; Ferguson et al., 2010; Sprent, 2007). Uninfected cells may intersperse the infected cells and nodule surfaces may contain prominent secondary aerenchyma called lenticels (Cermola et al., 2000; Tate et al., 1994). This nodule type is found in some of the major crop legumes, such as *P. vulgaris*, *Glycine max* (soybean) and *Arachis hypogea* (peanut).

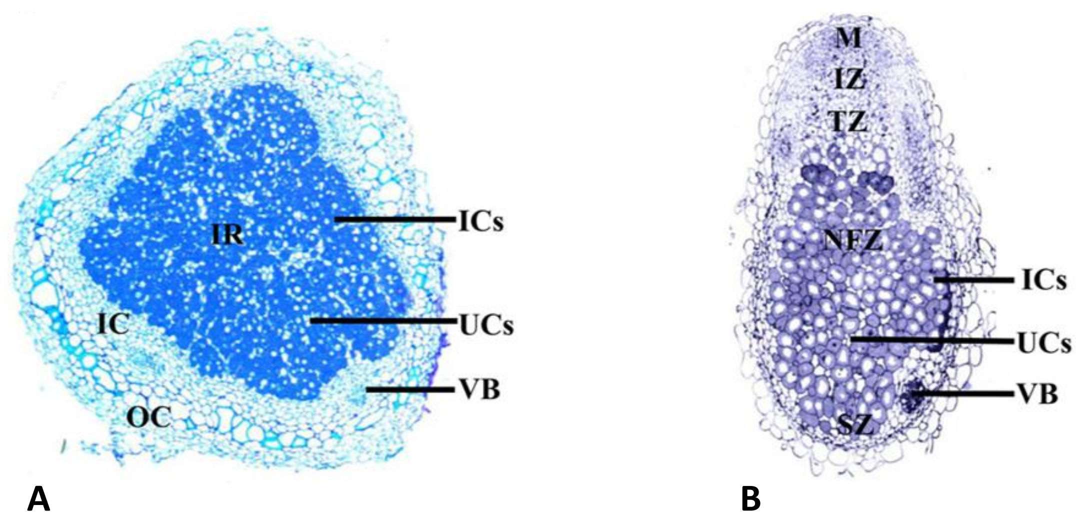


Figure 1.2: Images of (**A**), a cross section of a determinate nodule of *G. max* and (**B**), a longitudinal section of an indeterminate nodule of *M. trancatula*. Abbreviations: IR-infected region, ICs-infected cells, UCs-uninfected cells, IC-inner cortex, VB-vascular bundles, OC-outer cortex, M-meristem, IZ-invasion zone, TZ-transition zone, NFZ-N_2 fixation zone and SZ- zone of senescence. Image, without modification, from Brear et al. (2013).

1.2.2 Mechanism of N_2 fixation

The legume host fuels the energy demands of N_2 fixation and in return acquires fixed N. Carbon assimilated through photosynthesis, mainly in the form of sucrose, is transported into the root nodules (Kouchi & Yoneyama, 1986; Lodwig & Poole, 2010),

where, in the cytosol of nodule cells it is hydrolysed into glucose and fructose by sucrose synthase (Gordon et al., 1999). Fructose and glucose are oxidised to phosphoenolpyruvate (PEP) through glycolysis and PEP subsequently into dicarboxylates (Colebatch et al., 2004). Malate is the principle dicarboxylate transported via the Dct system into bacteroids where it undergoes assimilation by gluconeogenesis or is catabolized via enzymes of the tricarboxylic acid (TCA) cycle to provide the reductant and ATP required for N_2 fixation (Lodwig & Poole, 2010; Mus et al., 2016).

The nitrogenase complex, which catalyses N_2 fixation in bacteroids, consists of iron (Fe) and molybdenum-iron (MoFe) proteins. The Fe protein is a homodimer that contains a [4Fe-4S] cluster and 2MgATP binding sites while the MoFe protein is a heterotetramer containing a P-cluster [8Fe-7S] and a FeMo cofactor [7Fe-9S-Mo-C-homocitrate]. The FeMo cofactor is the site for substrate reduction (Burgess & Lowe, 1996; Lancaster et al., 2011; Seefeldt et al., 2004). Catalysis proceeds in three basic steps: (a) reduction of the Fe protein by electron donors such as ferredoxin, (b) transfer of electrons from Fe protein to the P-cluster and MgATP hydrolysis, and (c) transfer of electrons from P-cluster to the FeMo cofactor and subsequent N_2 reduction (Dixon & Kahn, 2004; Seefeldt et al., 2004). As nitrogenase is O_2 sensitive, oxidation is prevented by the host-controlled O_2 diffusion barrier operating across the nodule cell membranes, reversible O_2 binding by leghaemoglobin and high respiratory consumption of O_2 by bacteroids (Mus et al., 2016; Udvardi & Poole, 2013).

The NH_3 produced from N_2 fixation diffuses across the bacteroid membrane into the symbiosomal space where it is protonated into NH_4^+ and exported into the plant cell (Mus et al., 2016). NH_4^+ is assimilated by the plant using glutamine synthetase and glutamate synthase forming glutamine, which is the principal nitrogen metabolite for the synthesis of different amino acids, nucleic acids and other nitrogen-containing compounds (Mus et al., 2016). In legumes such as *P. vulgaris*, the nitrogen compounds are further metabolised into allantoin and allantoic acid via the purine biosynthetic pathway and exported to the rest of the plant through the xylem (Hungria & Kaschuk, 2014; White et al., 2007).

1.2.3 Effectiveness of N_2 fixation

Although the process of N_2 fixation proceeds through similar steps in most legumes (Section 1.2.1 and 1.2.2), the symbiotic relationships formed fix highly variable amounts

of N_2 (Buttery et al., 1997; Cardoso et al., 2012; Giller et al., 1998; Hungria et al., 2000; Hungria & Kaschuk, 2014). However, the outcomes of interactions can broadly be grouped into four categories (Howieson et al., 2005):

i. Non-infective: no nodulation occurs (Nod^-)
ii. Infective but ineffective: nodules form but no N_2 fixation occurs (Nod^+Fix^-)
iii. Partially effective (Nod^+Fix^+)
iv. Effective (Nod^+Fix^{++})

1.2.3.1 Causes of variability in N_2 fixation

In agriculture, symbiotically effective relationships (type iv) are desirable as they lead to the fixation of the most N, resulting in higher yields (Hungria et al., 2006). However, as not all symbioses are effective, an understanding of the causes of the variability is required, to form the basis for strategies to maximise the benefits from N_2 fixation. The outcomes of symbiotic relationships depend on legume × rhizobia × environment interactions.

Host

Legume symbioses have degrees of specificity mediated by the exchange of signalling molecules between the host and rhizobia (Section 1.2.1), and compatibility between the partners at each stage influences nodulation and N_2 fixation leading to the four outcomes (Oldroyd, 2013). As a generalisation, closely related legumes groups such as cultivars have similar N_2 fixation outcomes (i-iv above) with the same rhizobia, under controlled conditions. However, the amount of N_2 fixed may vary depending on differences in characteristics such as growth potential (N demand) and length of the vegetative period (Buttery et al., 1997; Graham et al., 2003; Hardarson & Atkins, 2003). Exceptions to the generalisation exist. In a study by Sadowsky and Cregan (1992), *Glycine max* cv. Hill was ineffective with *Bradyrhizobium elkanii* USDA 61, a strain that was effective on *G. max* cv. Williams. The ineffectiveness of USDA 61 on cv. Hill is linked to the Rj4 allele of the host (Faruque et al., 2015; Tsukui et al., 2013). The fact that these exceptions are not common suggests that in an agricultural context, cultivar choices may be important, but are not the greatest sources of variability in effectiveness of N_2 fixation.

Strain

The effectiveness of N_2 fixation in legumes varies depending on the rhizobial partner (Figure 1.3). To exploit this variability, numerous studies have successfully identified strains that achieve high rates of N_2 fixation with different legumes and in many cases used them as inoculants to improve yields (Bianco et al., 2013; Cardoso et al., 2012; Howieson et al., 2005; Hungria et al., 2000; Hungria et al., 2006; Rahmani et al., 2011; Waswa et al., 2014).

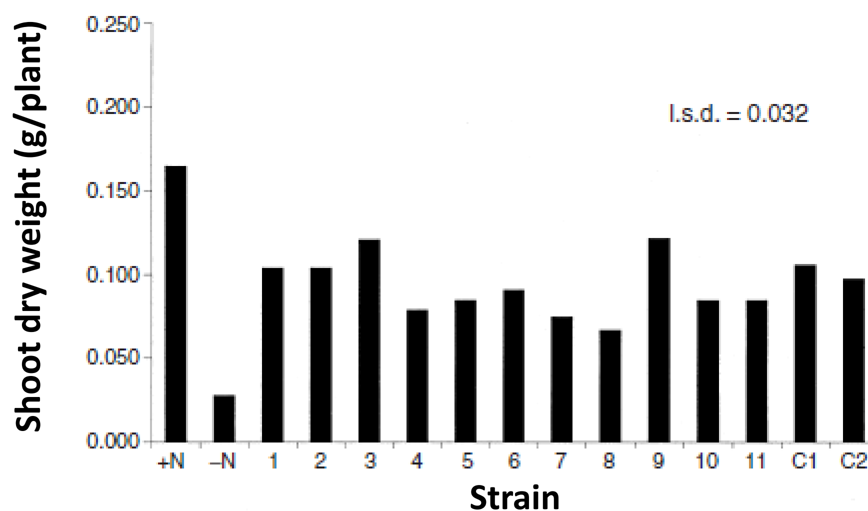


Figure 1.3: An example of variability in N_2 fixation in a single host due to differerent rhizobial strains. Data shows symbiotic effectiveness (measured as dry shoot weight per plant) of thirteen rhizobial strains on *Trifolium subterraneum* (clover) cv. Gosse (Collins et al., 2002).

Despite the selection and use of effective rhizobial strains as inoculants being a common practice for decades (Bullard et al., 2005), the basis for the observed differences in strain effectiveness is only beginning to be understood. Symbiotic interactions proceed through three major stages, of a) partner recognition, b) infection and invasion and c) release from infection threads, intracellular existence and N_2 fixation (Section 1.2.1 and 1.2.2). The variability in N_2 fixation (outcomes i-iv) is therefore expected to stem from differences in the success of strains at each of these stages.

At the beginning of the molecular dialogue, rhizobial strains need to have a *nodD* gene that can be activated by the legume-derived signalling molecules, such as flavonoids. Incompatibility at this stage leads to outcome (i) (Downie, 2014).

Following a successful signal recognition, the molecular dialogue persists during the infection and invasion phase (Niehaus et al., 1998). All the molecular determinants of a successful invasion are not known, but bacterial surface polysaccharides (e.g. lipopolysaccharides (LPS) and exopolysaccharides (EPS)) appear to play a significant role. In *M. trancatula,* an LPS mutant *Sinorhizobium meliloti* Rm6963 formed enlarged infection threads, and the outcome was an ineffective phenotype characterised by small white nodules (outcome ii). The authors hypothesized that LPS was necessary for continued division and persistence of the symbionts during infection and invasion, to suppress host defence mechanisms (Niehaus et al., 1998). In addition to LPSs, *nod* genes are required for the development of normal infection threads (Ardourel et al., 1994; Walker & Downie, 2000).

The molecular signalling during the internalisation of rhizobia into root cortical cells, differentiation into bacteroids, intracellular existence and N_2 fixation is very poorly understood, but again LPSs and extracellular polysaccharides appear to play a prominent role (Banba et al., 2001; Mathis et al., 2005). LPSs are modified as rhizobia develop into bacteroids (Kannenberg & Carlson, 2001) and as was shown with *Rhizobium* sp. strain NGR234, mutants unable to undertake this LPS modification, were partially effective (outcome iii) on siratro, *T. vogelii* and cowpea. They formed nodules with low levels of legheamoglobin and few bacteroid-containing cells (Broughton et al., 2006). The LPS modifications were hypothesised to mediate proximity and adhesion at the bacterium-plant interface, thereby affecting symbiotic functions that could include the invasion process, cell division within the symbiosomes, exchange of metabolites, and the suppression of the plant defence response (Kannenberg & Carlson, 2001). All these functions are important in determining the outcomes of symbioses.

Rhizobia also differ in the amounts of poly-β-hydroxybutyrate (PHB) (Lodwig et al., 2005; Trainer & Charles, 2006) they accumulate, and conflicting evidence exists on the effect of this trait on N_2 fixation. PHB is an intracellular carbon storage polymer produced by free-living rhizobia and bacteroids of determinately nodulating legumes (Lodwig et al., 2005; Trainer & Charles, 2006), although a recent study has shown indeterminate nodules can also accumulate small quantities of PHB (Terpolilli et al., 2016). PHB is speculated to be a means for rhizobia to hoard carbon to benefit subsequent generations (Ratcliff et al., 2008) or a sink for reductive power that allows the TCA cycle to operate under microaerobic conditions (Cevallos et al., 1996; Terpolilli

et al., 2016). PHB mutants are of lesser, equal or higher N_2 fixing capacity (Aneja et al., 2005; Cevallos et al., 1996; Lodwig et al., 2005) but the interpretation of findings from these studies is limited by the pleiotropic nature of PHB genes. Mutants are also defective in EPS production (Aneja et al., 2004; Trainer, 2009) and utilisation of various carbon sources (Cai et al., 2000; Cevallos et al., 1996; Lodwig et al., 2005). Additional studies that do not use mutants are required to further investigate whether strain differences in PHB accumulation are associated with variability in N_2 fixation.

Environment

Finally, abiotic and biotic environmental factors also influence the effectiveness of symbiotic interactions (Giller et al., 1998; Graham et al., 2003; Hungria & Vargas, 2000; Zahran, 1999). N_2 fixation is closely linked to the physiological state of the host, and environmental factors that reduce plant growth such as nutrient deficiencies, salinity, unfavourable pH, drought, diseases and extreme temperatures, influence the amount of N_2 fixed (Hungria & Vargas, 2000; Zahran, 1999). These stresses also impact on the survival of rhizobia and therefore their ability to colonise plant roots and form nodules (Graham et al., 1994; Zahran, 1999). N_2 fixation is also related to the plant demand for N, with soil N unavailability being conducive for N_2 fixation and conversely, high amounts of mineral N suppressing nodulation and N_2 fixation (Mortier et al., 2012; Zahran, 1999). Under field conditions, soil biota can influence N_2 fixation through microbial competition and antagonisms (Yates et al., 2011). For example, the presence of ineffective rhizobia reduces the overall effectiveness of symbioses as these form ineffective nodules that result in a reduction in the outputs of N_2 fixation (Gerding et al., 2014).

The above review of the literature shows that the host, rhizobia and the environment have roles in determining N_2 fixation outcomes. In an agricultural context, the choice of the host, at the cultivar level, sometimes influences N_2 fixation outcomes. However, cultivar × strain interactions are not reported in major crop legumes such as *P. vulgaris* (Buttery et al., 1997) and cultivar selection has not been the focal point of efforts to improve N_2 fixation, although attempts have been made to breed for N_2 fixation (Herridge & Rose, 2000). Additionally, cultivar choices in food crops are often under the influence of consumer preferences (Broughton et al., 2003) and therefore a slightly higher cumulative N_2 fixation is unlikely to make a compelling case for the adoption of a particular cultivar, unless other desirable attributes are already present or are successfully

bred into the cultivar. Consequently, a microbiological approach to enhancing N_2 fixation through the collection and screening rhizobia for N_2 fixation is more common (Bianco et al., 2013; Cardoso et al., 2012; Howieson et al., 2005; Hungria et al., 2000; Hungria et al., 2006; Rahmani et al., 2011; Waswa et al., 2014). The success with this approach is exemplified by the soybean story in Brazil where soybean inoculated with highly effective strains can fix over 300 kg N ha^{-1} Brazil (Hungria et al., 2006). Despite these successes, the benefits from inoculation with effective strains are sometimes limited by the presence of ineffective indigenous strains (Thies et al., 1991a). This challenge is greater in legumes that nodulate with a wide range of rhizobia (such as *P. vulgaris*) and to improve the yields of these legumes, in addition to symbiotic effectiveness, the challenge of rhizobial competition occasioned by their symbiotic promiscuity needs to be addressed.

1.2.4 Symbiotic promiscuity

Some legumes, such as *P. vulgaris* and *Macroptilium atropurpureum* (siratro), fix N_2 with different species of rhizobia (Martinez-Romero, 2003; Perret et al., 2000). This characteristic, known as promiscuity, is in contrast to that of legumes such as *P. sativum* and *Vicia faba* (broad bean) that have a restricted host range (Perret et al., 2000). Symbiotic promiscuity increases the number of potential competitors to inoculant strains, leading to a high incidence of inoculation failure (Vlassak et al., 1997). Considerable research efforts have gone into understanding the molecular basis of symbiotic promiscuity (Perret et al., 2000) and known mechanisms include the synthesis of diverse plant signalling molecules, multiple *nodD* alleles in rhizobia and perception of diverse Nod factors by the host (Del Cerro et al., 2015; Laeremans & Vanderleyden, 1998).

In *P. vulgaris,* roots exude a cocktail of flavonoids that interact with NodD proteins and activate transcription of *nod* genes that code for the enzymes that synthesize Nod factors (Hungria et al., 1992; Hungria et al., 1991; Laeremans & Vanderleyden, 1998). Since different flavonoids show specificity in their ability to interact with various *nodD* genes, the various flavonoids produced by *P. vulgaris* allow the initiation of a molecular dialogue with diverse rhizobia (Laeremans & Vanderleyden, 1998; Spaink et al., 1987). In addition to roots exuding diverse flavonoids, rhizobia that nodulate *P. vulgaris* commonly harbour multiple copies of *nodD* genes. As an example, *Rhizobium tropici* CIAT 899 has five *nodD* alleles (Del Cerro et al., 2015; Ormeño-Orrillo et al., 2012). Having different

nodD alleles extends the host range of rhizobia as it increases possible *nodD*-flavonoid combinations (Peck et al., 2006).

Secondly, *P. vulgaris* roots can perceive and respond to different Nod factors (NFs), further enabling promiscuity. NFs are a family of lipo-chitooligosaccharides of species-specific size and saturation (Long, 1996; Perret et al., 2000). Rhizobia-specific substitutions of the reducing and non-reducing glucosamine residues occur (coded by several *nod* genes) and, in addition to the NF core structure (coded by *nodABC*), are known determinants of host range through specific interactions membrane-bound NF receptors on hair root cells (Perret et al., 2000; Radutoiu et al., 2003; Radutoiu et al., 2007; Rodpothong et al., 2009). Although some strains of rhizobia that nodulate *P. vulgaris* produce few NFs (Poupot et al., 1993) other strains produce an incredibly high number of different NFs with up to 52 types reported from *Rhizobium tropici* CIAT 899 (Estévez et al., 2009; Guasch-Vidal et al., 2013; Morón et al., 2005). This variability of NFs produced by strains that can nodulate *P. vulgaris* suggests that the host can perceive a wide range of NFs.

The above mechanisms and others (Perret et al., 2000) enable symbiotic promiscuity, but actual promiscuity requires the presence of a promiscuous host and multiple microsymbiont partners. In the absence of diverse microsymbionts, the capacity for promiscuity does not affect inoculation outcomes. However, in the presence of multiple compatible rhizobia, competition occurs between the native rhizobia and inoculant strain for the formation of nodules (Vlassak et al., 1997). Rhizobial groups differ in characteristics such as growth rates and utilization of carbon compounds, traits that might influence competition outcomes and knowledge on the predominance of taxa or strains is expected to be useful to the management of rhizobial competition. Such knowledge gained through taxonomic and diversity studies, also informs on the genetic relationships among rhizobia, rhizobial preferences of the host and the dynamics of exchange of genetic material (Martinez-Romero, 2003).

1.3 Methods to study taxonomy and diversity of rhizobia

Various tests help determine whether a strain belongs to a known or novel taxon. These tests are guided by the Bacteriological Code, subsequent revisions, directions from the International Journal of Systematic and Evolutionary Microbiology, or minimal standards set by International Code of Nomenclature of Prokaryotes subcommittees

(Graham et al., 1991; Lapage et al., 1992; Parker et al., 2015; Tindall et al., 2010). Once a primary test or tests place a strain unequivocally in a described taxon, further characterization for purposes of identification is not necessary (Tindall et al., 2010).

Rhizobial phenotypic characteristics such as cell shape, colony morphology, pH range, temperature range, biochemical tests, and chemotaxonomic characters such as fatty acid profiles are commonly used in the identification of rhizobia (Tindall et al., 2010). However, as phenotypic traits on their own cannot be used for unequivocal assignment to a taxon, these are mainly used as pre-requisites to DNA methods or to corroborate DNA methods.

The most frequently used technique for the identification of rhizobia is the sequencing of 16S rRNA genes. Ribosomal ribonucleic acid (rRNA) genes (16S, 23S and 5S) are highly conserved because of the fundamental role of ribosomes in protein synthesis and carry considerable evolutionary information (Woese, 1987). Although 23S rRNA genes carry the most genetic information, their large size makes them unsuitable for routine sequencing and, consequently, the smaller-sized 16S rRNA genes (~1,500 bp) are the primary genes for phylogenetic identification (Rosselló-Móra & Amann, 2001; Tindall et al., 2010). In general, strains with less than 97% sequence similarity belong to different species while those with a similarity less than 95% belong to different genera (Gevers et al., 2005; Stackebrandt & Goebel, 1994). Partial sequences of 16S can be used for phylogenetic purposes, but at least 1300 bp are required for a reliable grouping (Yarza et al., 2014).

Due to the conserved nature of 16S rRNA genes, they are useful in identifying a strain to the genus level but may fail to differentiate between closely related species (Fox et al., 1992). For example, the type strain for the recently described *Rhizobium sophorae* strain CCBAU 03386 shares 100% 16S rRNA sequence similarity with *Rhizobium laguerreae* FB206^T but showed only 25.8% relatedness by DNA-DNA hybridization (DDH) (Jiao et al., 2015). An additional drawback with the use of 16S rRNA is that some bacteria have multiple copies of 16S rRNA genes that may show sequence divergence. These copies are often identical (Boucher et al., 2004; Klappenbach et al., 2001) but may vary by up to 6% (Wang et al., 1997). Importantly, as with other genes, 16S rRNA may also be subject to recombination or horizontal gene transfer further reducing its reliability when used as the sole gene for phylogenetic studies (Gevers et al., 2005).

These drawbacks necessitate the use of additional genes that increase the correlation of data obtained with DDH. Increasingly, in addition to 16S rRNA comparison, strains are analysed by comparing several housekeeping genes in a multilocus sequence analysis (MLSA). The number of housekeeping genes used in an MLSA for rhizobia species identification varies and examples includes two by Vinuesa et al. (2005), three by Aserse et al. (2012), five by Rivas et al. (2009) and six by Degefu et al. (2013). The sequences of the housekeeping genes are either compared individually or more commonly as concatenated sequences (Tindall et al., 2010). MLSA is used to provide deeper branching in phylogenetic trees to differentiate species closely related by 16S rRNA sequences.

In summary, to correctly assign a root nodule bacterium to a known taxon, it is required that after confirmed and reproducible nodulation of a legume (Graham et al., 1991), the strain should be assessed by method(s) empirically proven to correlate with DNA-DNA hybridization (e.g. 16S rRNA for genus assignment and MLSA for species assignment). For the description of novel taxa, DDH and other supporting discriminant tests prescribed by the Bacteriological Code are required to be undertaken (Tindall 2010).

Classifications below the species level include biovars (bv.), symbiovars (sv.) and strains. Generally, in bacteriology, a biovar refers to a group of strains discernible from others in a species by physiological or biochemical means. However, when applied to rhizobia, it is indicative of host range, and the term symbiovar (symbiotic variety) has been suggested to be more suitable as it more accurately reflects symbiotic genes and host range (Rogel et al., 2011). As an example, *R. leguminosarum* has symbiovars phaseoli, trifolii and viciae that enable the species to nodulate *P. vulgaris*, *Trifolium* spp. and *Vicia* spp. respectively (Rogel et al., 2011). Symbiovars are assigned based on host range studies or by DNA sequencing of *nod* genes such as *nodC* or *nodA* (Faghire et al., 2012; Mnasri et al., 2012; Rogel et al., 2011; Rouhrazi et al., 2016). Diversity at the strain level is mainly assessed using DNA fingerprinting techniques. These methods include enterobacterial repetitive intergeneric consensus (ERIC) (De Bruijn, 1992) and randomly amplified polymorphic DNA-PCR (Richardson et al., 1995). These techniques have the power to distinguish rhizobia at the strain level(Thies et al., 2001).

1.4 *Phaseolus vulgaris*

1.4.1 Cultivation of *P. vulgaris*

P. vulgaris is an annual, mostly self-pollinated leguminous plant cultivated for food in many parts of the world. *P. vulgaris* originated in Mesoamerica and diverged, about 110,000 - 165,000 years ago, into Mesoamerican and Andean gene pools that are now partially reproductively isolated (Bitocchi et al., 2012; Koinange & Gepts, 1992; Mamidi et al., 2013; Schmutz et al., 2014). *P. vulgaris* was then domesticated independently in Mesoamerica and the Andes approximately 7,000 years ago, and cultivation has over the centuries spread from these two domestication centres to large areas in the tropics and subtropics.

Through a mixing of the two gene pools during the many years of cultivation, with the rare introgression between domesticated and wild relatives, highly variable landraces and cultivars are grown today (Broughton et al., 2003). One such variability is in growth habit. Determinate (bush) types are marked by a terminal inflorescence during developmental stages and remain erect and of short stature. Indeterminate types have long vines, with growth continuing until plant senescence. Indeterminate types are either erect, prostrate or climbers, and have higher grain yields than determinate types (Kelly et al., 1987).

Presently, *P. vulgaris* is the most important crop legume for direct human consumption with over 45 million tonnes of beans produced in 2013 (FAOSTAT), primarily in Latin America, Sub-Saharan Africa (SSA), and Asia (Figure 1.4). The consumption of *P. vulgaris* is mainly as dry grain (dry beans) or as a green pod (green beans). Less frequently, the immature seed or leaves are eaten. Beans provide dietary protein, carbohydrates, fibre and are a good source of vitamins and minerals such as potassium and phosphorous (Broughton et al., 2003). Consumption varies across the globe but is highest in Sub-Saharan Africa, with the annual per capita consumption as high as 66 kg in parts of Kenya (Broughton et al., 2003).

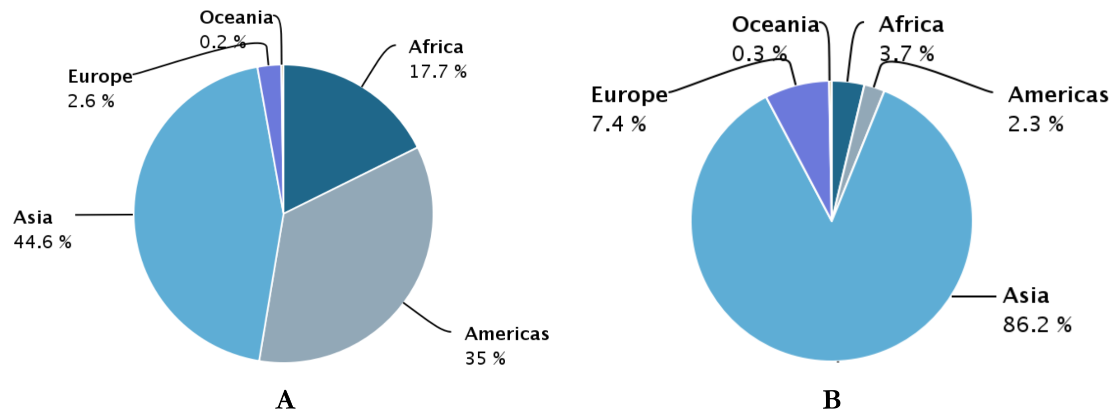


Figure 1.4: Average production share of (**A**) dry beans and (**B)** green beans by region for the period between 1993 and 2013 (FAOSTAT).

Production of *P. vulgaris* is mainly in small-scale farm holdings ranging from 1-10 ha (Broughton et al., 2003; Graham & Ranalli, 1997). Yields are highly variable. For example, in 2013, average dry bean yields of 4,907 kg ha^{-1}, 3183 kg ha^{-1}, and 2093 kg ha^{-1} were harvested in Ireland, Netherlands and the USA, respectively. In the same year, average yields were 659 kg ha^{-1} in Kenya (FAOSTAT). Reasons for the differences in yield include differences in the yield potential of cultivars grown, climatic conditions, and low soil fertility in critical nutrients such as N (Broughton et al., 2003; Graham & Ranalli, 1997; Katungi et al., 2009; Marinus, 2015).

P. vulgaris is capable of fixing N_2, and although the potential for *P. vulgaris* to fix N_2 is reputed to be low (Bliss, 1993; Peoples et al., 2009), the use of effective, competitive and well adapted rhizobial strains in soils with low N can lead to high yields. In Brazil, the use of rhizobial inoculants was reported to increase yields of *P. vulgaris* by up to 900 kg ha^{-1} (Hungria et al., 2000).

1.4.2 Effectiveness of *P. vulgaris* symbiosis

Of the agriculturally important legumes, *P. vulgaris* shows relatively low rates of N_2 fixation (Herridge et al., 2008; Peoples et al., 2009). The global average shoot N fixed in farmers' fields by *P. vulgaris* is 15 kg N ha^{-1} (Peoples et al., 2009). This rate is well below soybean (137 kg N ha^{-1}), pea (108 kg N ha^{-1}), lentil (71 kg N ha^{-1}), cowpea (20 kg N ha^{-1}), pigeon pea (59 kg N ha^{-1}) and groundnut (103 kg N ha^{-1}) (Peoples et al., 2009).

The comparatively low rates of N_2 fixation in *P. vulgaris* observed in the field are attributed to several factors. The first of these are inherent characteristics that include a

short vegetative period (Graham et al., 2003), and a low respiratory efficiency of its N_2 fixation in comparison to other legumes (Witty et al., 1983). Secondly, low N_2 fixation rates result from the cultivation of a plant with relatively high nutrient requirements in low input systems deficient in nutrients critical for growth and N_2 fixation such as phosphorus and molybdenum (Bressers, 2014; Brodrick et al., 1995; Broughton et al., 2003; Giller et al., 1998).

Furthermore, genetically heterogeneous cultivars of *P. vulgaris* are grown, and these exhibit differences in the amounts of N_2 they fix (Bliss, 1993; Buttery et al., 1997; Hardarson & Atkins, 2003), symbiotic tolerances to nitrate (Park & Buttery, 1989), nutrient use efficiencies (Ramírez et al., 2013; Vadez et al., 1999) and affinities towards various rhizobia (Aguilar et al., 2004; Oliveira et al., 2011). In spite of all this variability, cultivar choice in many areas is primarily based on consumer taste preferences and less on traits related to N_2 fixation (Broughton et al., 2003; Graham & Ranalli, 1997).

Another factor that leads to low level of N_2 fixation in *P. vulgaris* is a low degree of inoculant use (Peoples et al., 2009), and the use of a small number of inoculant strains in very diverse ecological zones. For example, *R. tropici* CIAT 899 is widely used in eastern and southern Africa to inoculate *P. vulgaris* (Bala et al., 2011) despite the strain being poorly adapted to some soils in the region (Anyango et al., 1998). Additionally, the promiscuity of *P. vulgaris* exacerbates the problem of competition. Most soils used to cultivate *P. vulgaris* contain high numbers of variably effective rhizobia compatible with *P. vulgaris,* and these resident rhizobia are known to inhibit responses to inoculation (Anyango et al., 1995; Hungria & Vargas, 2000; Kawaka et al., 2014; Thies et al., 1991a)

1.4.3 Diversity of *P. vulgaris*-rhizobia

P. vulgaris nodulates and fixes N_2 with a broad range of rhizobial species in up to five genera (Table 1.1). The rhizobial species listed nodulate the host *in-situ,* were trapped from soils in the glasshouse, or nodulated following inoculation as pure cultures (*ex-situ*). Generally, strains that nodulate *P. vulgaris in-situ* belong to the genus *Rhizobium,* but this range is extended considerably to include species in other genera when *ex-situ* nodulators are considered (Table 1.1). A narrower range of rhizobia is isolated from the field than from relocated soil samples or from inoculation with serially diluted soil (Alberton et al., 2006; Bala et al., 2001) and one explanation for this is that *P. vulgaris* may have a level of

selectivity for certain rhizobial types in the field. Alternatively, *ex-situ* or *in-situ* methods of isolating rhizobia may favour the isolation of certain rhizobial species.

In addition to the range of species capable of fixing N_2 with the host, a greater number of rhizobial species can elicit ineffective nodules on *P. vulgaris* (Martinez-Romero, 2003; Michiels et al., 1998). This is because the stringency of Nod factor induction of nodule organogenesis has a lower stringency than is required for bacterial infection (Oldroyd & Downie, 2008). Rhizobia reported to form ineffective nodules with *P. vulgaris* are not included in Table 1.1.

1.4.4 Spread of *P. vulgaris* rhizobia across the globe

Different rhizobia nodulate, and fix N_2 with *P. vulgari*s in various regions of the world (Table 1.1). The genetic diversity of these microsymbionts appears to increase with distance away from the Mesoamerican and Andean regions, the centres of origin of *P. vulgaris*, with only six species listed in Table 1.1 described from the two areas. This observation is consistent with the distance-decay theory of biological similarity where the diversity of organisms increases with distance from a source depending on dispersal limitation and niche difference (niche characteristics and competition) (Bell, 2010; Martiny et al., 2011; Nekola & White, 1999).

Generally, at least four factors influence the nature, distribution, and population densities of *P. vulgaris* rhizobia across the globe. The first factor is the presence or absence of a legume host. This host could be *P. vulgaris* or an alternative compatible host. Rhizobia may have co-evolved with legumes (Aguilar et al., 2004; Heath et al., 2012; Lowther & Patrick, 1993; Sprent, 2007) and nodulation is believed to have started soon after the origin of legumes 60 million years ago (Lavin et al., 2005; Sprent, 2008). Legumes such as *P. vulgaris* have been reported to preferentially nodulate with certain rhizobial types (Aguilar et al., 2004; Caballero-Mellado & Martinez-Romero, 1999; Montealegre et al., 1995) and can therefore potentially shape rhizobial diversity in soil.

Table 1.1: List of rhizobial species known to contain strains that nodulate and fix N_2 with *P. vulgaris*

	Species	Type strain†	Nodulation			Origin of strain	Reference
			In-situ	*Ex-situ*			
			Field	Soil	Culture		
1	*R. tropici*	✓	✓			Colombia	Martinez-Romero et al. (1991)
2	*R. etli*	✓	✓			Mexico	Segovia et al. (1993)
3	*R. gallicum*	✓	✓			France	Armager et al. (1997)
4	*R. giardnii*	✓	✓			France	Armager et al. (1997)
5	*R. lusitanum*	✓		✓		Portugal	Valverde et al. (2006)
6	*R. phaseoli*	✓				NS*	Ramirez-Bahena et al. (2008)
7	*R. vallis*	✓				China	Wang et al. (2011)
8	*R. freirei*	✓		✓		Brazil	Dall'Agnol et al. (2013)
9	*R. paranaense*	✓		✓		Brazil	Dall'Agnol et al. (2014)
10	*R. azibense*	✓	✓			Tunisia	Mnasri et al. (2014)
11	*R. ecuadorense*	✓	✓			Ecuador	Ribeiro et al. (2015)
12	*R. acidisoli*	✓				Mexico	Román-Ponce et al. (2016)
13	*R. mesoamericanum*	✓		✓		Mexico	López-López et al. (2012)
14	*R. leucaenae*	✓				Brazil	Ribeiro et al. (2012)
15	*R. grahamii*	×		✓		Mexico	López-López et al. (2012)
16	*Sinorhizobium americanum*	×		✓		Tunisia	Mnasri et al. (2012)
17	*Burkholderia phymatum*	×		✓		Morocco	Talbi et al. (2010)
18	*R. sophorae*	×			✓	China	Jiao et al. (2015)
19	*R. sophoriradicis*	×			✓	China	Jiao et al. (2015)
20	*R. pisi*	×			✓	NS	Ramirez-Bahena et al. (2008)
21	*R. leguminosarum*	×			✓	NS	Ramirez-Bahena et al. (2008)
22	*R. mongolense*	×			✓	Mongolia	Van Berkum et al. (1998)
23	*Parabulkholderia nodosa*	×		✓		Brazil	Dall'Agnol et al. (2016)
24	*Sinorhizobium fredii*	×		✓		Spain	Herrera-Cervera et al. (1999)
25	*Sinorhizobium meliloti*	×		✓		Spain	Zurdo-Piñeiro et al. (2009)
26	*Bradyrhizobium elkanii*	×			✓	NS	Hungria and Kaschuk (2014)

†Type strain isolated from *P. vulgaris*
*NS- Not specified

The methods and rates of dispersal of rhizobia from one location to another also determine prevalence of strains at a locality. Dispersal within short geographical distances (e.g. metres) is dependent on factors such as wind, slope and surface runoff (Lowther & Patrick, 1993; Woomer, 1990) while dispersal over long distances, that may span continents, is by the intentional or unintentional transport by humans. For example, rhizobia can be carried on or in legume seeds where they can remain viable for an extended period (Mora et al., 2014; Pèrez-Ramirez et al., 1998). Consequently, the movement of legume seeds across markets unintentionally aids in the dispersal of rhizobia across geographical locations. *R. tropici*, widely distributed in Brazilian soils, is believed to have been introduced into Brazil from the Andes together with *P. vulgaris* seed by immigration and trade (Gomes et al., 2015). Rhizobial inoculant strains are also commonly sold or exchanged across countries with, for example, *Rhizobium tropici* CIAT 899, isolated from Colombia (Martinez-Romero et al., 1991) being distributed as an inoculant for *P. vulgaris* in Kenya (Bala et al., 2011).

Upon dispersal, the continued presence of strains depends on the adaptability of the dispersed strains to the environmental conditions in their new habitats. Differences in intrinsic characteristics of the strains such as tolerance to low or high pH, soil temperature, soil moisture, salinity, soil texture and nutrient efficiencies play a role in determining survival (Anyango et al., 1995; O'Hara, 2001; Zahran, 1999). Additionally, the availability of a compatible host will influence survival as the levels of rhizobial populations are significantly enhanced in the presence of a compatible host (Vlassak et al., 1996).

Lastly, genetic exchange between rhizobia shapes the diversity of microsymbionts in any given region. Although the major reports on the genetic exchange between rhizobia mainly relate to *Mesorhizobium* sp. (Nandasena et al., 2007; Sullivan et al., 1995), genetic exchange exists in the many rhizobial species that nodulate *P. vulgaris*. The first set of evidence for this genetic exchange can be inferred from phylogenetic studies. In the genus *Rhizobium*, species with relatively different chromosomal genetic makeup (<70%) may have identical symbiotic genes. For example, highly conserved symbiotic genes, symbiovar *phaseoli*, enable species such as *R. etli*, *R. phaseoli*, *R. leguminosarum*, *R. sophorae* and *R. sophoriradicis* to nodulate *P. vulgaris* (Jiao et al., 2015; Rogel et al., 2011; Rouhrazi et al., 2016). The second set of evidence is experimental with for example Rao et al. (1994) and Brom et al. (2000) able to demonstrate the transmissibility of symbiotic plasmids

from *R. leguminosarum* bv. trifolii and *Rhizobium etli* respectively into non-nodulating strains and consequently transferring nodulation ability. With the continuous introduction of rhizobia into new areas, previously non-symbiotic rhizobia may acquire symbiotic genes that may allow them to nodulate legumes.

1.4.5 Rhizobia competition in *P. vulgaris* symbiosis

The promiscuity of *P. vulgaris* and the widespread dispersal of compatible rhizobia make rhizobial competition an important aspect of *P. vulgaris* symbiosis. Competition for nodulation may arise when more than one rhizobial genotype capable of infecting a legume is within the plant's rhizosphere (Yates et al., 2011). In *P. vulgaris,* this competition occurs between species of different genera such as *S. americanum* and *R. gallicum* in Tunisian soils (Mnasri et al., 2012); species of the same genus e.g. *R. etli*, *R. gallicum*, *R. giardnii*, *R. leguminosarum* and *S. fredii* in a soil in Spain (Herrera-Cervera et al., 1999); or between strains belonging to a single species (Pinto et al., 2007).

Rhizobial competition has considerable influence on the outcomes of *P. vulgaris* inoculation as it can present a barrier to the ability of an inoculant strain to occupy a significant portion of the host nodules (Thies et al., 1991a). Thies et al. (1991b), using 29 legume-site combinations estimated that 59% of the yield variability following inoculation of legumes was due to competition between the inoculant strain and the native rhizobia. Plants can only form a limited number of nodules, autoregulated through several mechanisms that include the use of a shoot-derived signal (Mortier et al., 2012). Therefore, the preclusion of an elite strain from some or all the nodules on a legume, by ineffective strain(s), leads to a reduction in the overall amount of N_2 fixed (Denton et al., 2002; Gerding et al., 2014; Thies et al., 1991a).

1.4.5.1 Determinants of rhizobial competition outcomes

Considerable research efforts have gone into trying to better understand rhizobia competition (Thies et al., 1991a; Triplett & Sadowsky, 1992; Yates et al., 2011). The legume host has a role in determining the outcomes of competition as exemplified in work by Sadowsky and Cregan (1992). In the study, *Glycine max* cv. Hill carrying the Rj4 allele was shown to nodulate poorly with *Bradyrhizobium elkanii* strains USDA 61 and USDA 438. In contrast, *Glycine max* cv. Hill nodulated well with *B. elkanii* USDA 110. The basis of the incompatibility remains poorly understood but is suspected to stem from a discordancy between the hosts carrying the Rj alleles and the T3SS effectors

produced by the strains (Faruque et al., 2015; Tsukui et al., 2013). In *P. vulgaris*, the observation that wild accessions of *P. vulgaris* preferentially nodulate with rhizobia from their geographical origin (Aguilar et al., 2004) and that cultivar RAB39 preferentially nodulates with *R. tropici* UMR1899 (CIAT 899) in the presence of *R. etli* (Montealegre et al., 1995) further supports the idea that hosts influence the outcomes of rhizobial competition. More recently in clover, Yates et al. (2008), demonstrated selective nodulation in two geographically and phenologically distinct clovers. *Trifolium purpureum* nodulated with its effective strain WSM1325, despite WSM1325 being outnumbered 100-fold by the ineffective WSM2304. Similarly, *T. polymorphum* preferentially nodulated with its effective strain WSM2304 even when WSM2304 was outnumbered 100-fold by the ineffective WSM1325.

In regards to the microsymbionts, investigations have centred on the impact of rhizobial population densities in soils on the outcomes of inoculation and the consensus appears to be that population density of soil rhizobia is inversely related to the inoculation response (Singleton & Tavares, 1986; Thies et al., 1991a, 1991b; Weaver & Frederick, 1974). However, the specific numbers reported hampering inoculation response varies with the location. For example, Thies et al. (1991a) reported that 93 rhizobia per g of soil inhibited a response to the inoculation of *P. vulgaris* in Kuiaha, Hawaii, in soils with growth-limiting amounts of N. In a study by Vargas et al. (2000), 700 rhizobia per gram of soil hindered the successful inoculation of *P. vulgaris* in the Cerrados of Brazil. In contrast, responses to the inoculation of *P. vulgaris* have been reported in soils with 10^3 (Vlassak et al., 1996) and 10^5 rhizobia per g of soil (Hungria et al., 2000).

Inoculation responses in the presence of high populations of background rhizobia (Hungria et al., 2000; Vlassak et al., 1996) are an intriguing phenomenon. For a response to inoculation to occur in these soils, the inoculant strain would need to out-compete the large populations of indigenous rhizobia for nodule occupancy. This would require the inoculant strain to overcome the positional advantage that the soil rhizobia have in regards to accessing the infection sites of the developing root systems (López-García et al., 2002). *P. vulgaris* is often cultivated in soils containing 10^4-10^6 rhizobia g^{-1} of soil (Alberton et al., 2006; Andrade et al., 1999; Anyango et al., 1995; Hungria et al., 2000; Kawaka et al., 2014; Langwerden, 2014) and how a response to inoculation in such high rhizobial backgrounds can occur is unknown.

The phenomenon above may be related to at least four factors: (a) the inoculum rate, (b) the ability of rhizobial strains involved to multiply in the rhizosphere of the host, (c) a host-mediated preferential nodulation with the inoculant strain or (d) an environmental influence on the interactions.

The primary mode of inoculant delivery globally is through the application of rhizobia in peat directly onto seeds, with the aid of different adhesives (Herridge et al., 2014). With this delivery method, the rhizobia carrying capacity of the peat and the size of the seed, limit the number of rhizobia applied. The recommended rhizobial rates per seed are 10^3, 10^4 and 10^5 for small, medium and large seeded legumes respectively (Catroux et al., 2001b; Hungria et al., 2005; Lupwayi et al., 2000). The inoculum rates are unlikely to be the reason for different outcomes in researcher-managed experiments of Thies et al. (1991a) (93 cells g^{-1} soil inhibiting inoculation response) and Vlassak et al. (1996) (inoculation response observed in the presence of 10^3 cells per g^{-1} soil). However, variability in inoculum rates in farmer-managed fields, mainly stemming from differences in quality of the inoculant (Balume et al., 2015) may impact inoculation outcomes. In soils with low populations of native rhizobia, or in soils with no compatible rhizobia, increasing the dosage of the inoculant strain results in increases in the inoculation response (Hume & Blair, 1992; Patrick & Lowther, 1995). However, the effect of inoculation density on nodule occupancies in *P. vulgaris,* in the presence of high populations of resident rhizobia, is poorly understood.

Another factor that may be related to an inoculation response in soils with a high population of native rhizobia is a superior rhizosphere competence of the inoculant strain in comparison to the native strains. Differences in rhizosphere competence may arise from differences in chemotaxis, motility, biofilm formation (Cooper, 2007; Frederix et al., 2014), utilisation of certain root exudates (Streit et al., 1992; Wielbo et al., 2007), and growth rates (Li & Alexander, 1986). Inoculations in two successive years were required for an inoculation response in the study by Hungria et al. (2000) suggesting a build-up of the inoculant strain in the soil may have facilitated the response to inoculation. However, other factors such as a reduction the levels of the soil N may also have been responsible for the response to inoculation in the second year. It is currently not known whether rhizosphere competence could help inoculant strains overcome a naturalised rhizobia population in the range of 10^4-10^6 cells per g of soil.

The environment also influences the outcomes of competition. The role of biotic and abiotic factors in the growth, survival and symbiotic performance of rhizobia is well documented (Vlassak et al., 1997; Zahran, 1999), but uncertainties still exist. One such uncertainty is the role of soil mineral N in determining rhizobial competition outcomes. In *P. vulgaris*, low levels of soil N can stimulate early plant growth and consequently enhance N_2 fixation (Da Silva et al., 1993; Müller et al., 1993; Tsai et al., 1993). However, higher N rates do not provide the "starter effect" and lead to a decrease in N_2-fixation by suppressing nodule numbers and weights (Da Silva et al., 1993; Leidi & Rodriguez-Navarro, 2000; Tsai et al., 1993). The mechanisms of nitrate suppression of nodulation and N_2 fixation are not fully known but may involve similar mechanisms as those participating in the autoregulation of nodulation (Mortier et al., 2012; Reid et al., 2011) or through decreased supply of photoassimilates to nodules (Fujikake et al., 2003).

While the roles of soil N in enhancing or suppressing nodulation and N_2 fixation are well documented, the influence of N levels on rhizobial competition and nodule occupancy is less clear. In studies by Vargas et al. (2000) and Caballero-Mellado and Martinez-Romero (1999), soil N was observed to influence the rhizobial types nodulating *P. vulgaris*. For example, nodule occupancy by *R. tropici* CIAT 899 decreased from 91% to 18% with the increase in levels of soil N while simultaneously, nodule occupancy by the native strains rose from 9% to 82% with the increase in soil N (Vargas et al., 2000). The authors hypothesised that the alteration in nodule occupancies resulted from either difference in nitrate tolerance by the strains or the production of altered root exudates that interact with the symbiosis and influence nodule occupancy outcomes (Caballero-Mellado & Martinez-Romero, 1999; Vargas et al., 2000). The effect of N is, however, equivocal as other reports indicate soil N has no influence on rhizobial competition and nodule occupancy outcomes (Abaidoo et al., 1990; Glyan'ko et al., 2009). As *P. vulgaris* is grown in farms with high soil N heterogeneity (Vanlauwe et al., 2016), the influence of N on the competition outcomes of rhizobia that nodulate this host needs to be clarified.

From the review, above, many gaps in our current knowledge of the determinants of rhizobial competition outcomes are apparent. These areas need to be investigated. However, any advances in knowledge in this important area of competition require tools to assist in determining rhizobia in nodules, in the rhizosphere and in the bulk soil.

1.4.5.2 Strain typing in competition studies

Rhizobial competition studies require methods to identify nodule occupants and to assess population densities. These methods should allow (a) reliable discrimination of rhizobia, preferably at the inter-strain level (b) rapid identification (c) cost-efficient analyses and (e) amenability to use with many strains. Currently, several methods are available, and these are variably suited to different studies. Methods in use exploit the phenotypic, proteomic or genotypic differences among strains.

Phenotypic, proteomic and genotypic methods

Phenotypic techniques include assessment of colony morphologies and biochemical characteristics. Melanin production (Blanco et al., 2010; Castro et al., 2000) and ELISA (Castro et al., 2000; Spriggs & Dakora, 2009) are examples of phenotypic characteristics that have been used to study rhizobia competition. Phenotypic characters do not offer sufficient discrimination between closely related strains, and in some instances, distantly related strains exhibit similar characteristics. For example, Cardoso et al. (2012) showed strains 2608 and 1281 were virtually indistinguishable using 12 phenotypic characters that included melanin production, bromothymol blue (BTB) reaction and a battery of colony characteristics. However, using nearly full length 16 S rRNA sequences, 2608 was found to be related to *Rhizobium giardnii* strains, while 1281 grouped with members of the genus *Stenotrophomonas.* With ELISA, in addition to requiring sampling of nodules, cross reactions with other rhizobia can lead to false positive results (Hungria et al., 2000; Spriggs & Dakora, 2009).

Proteomic approaches, such as the matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) can differentiate between species, including very closely related species of *Rhizobium* indistinguishable by 16S (Ferreira et al., 2011; Martínez-Molina et al., 2016), between biovars (Šalplachta et al., 2013) and between some sub-specific groups (López Díez et al., 2016). However, the discriminatory power of MALDI-TOF MS below the species level is disputed (Karlsson et al., 2015), and therefore the utility of this method for strain level identification and subsequently competition studies is not clear.

Genotyping methods include those based on DNA banding patterns, and may or may not involve a PCR amplification stage as well as DNA sequencing methods. Genotypic methods commonly used in rhizobia competition studies include DNA fingerprinting methods like enterobacterial repetitive intergeneric consensus (ERIC)-PCR, repetitive

extragenic palindromic (REP)-PCR, GTG_5-PCR and PCR with directed primers such as RP01 (Denton et al., 2002; Gerding et al., 2014).

Among other drawbacks, a general one shared by phenotypic, proteomic and genotypic methods is the inability to rapidly and simultaneously identify nodule occupants on nodules on an entire root system. This problem is magnified substantially in legumes such as *P. vulgaris* that can form hundreds of nodules on a single root system. The sampling of nodules or isolates from nodules (Spriggs & Dakora, 2009; Wongphatcharachai et al., 2015), reduces the robustness of the data collected. Additionally, with these methods, the enumeration of bacterial numbers on roots, rhizosphere and in the bulk soil is challenging.

Marker genes

The insertion of genes that specify a selectable or phenotypically identifiable characteristic has for decades been useful in the ecological study of bacteria. Markers used in rhizobial competition studies include *gusA* (Denton et al., 2003; Pitkäjärvi et al., 2003; Reeve et al., 1999; Shamseldin, 2007; Wilson et al., 1995), *lacZ* (Krishnan & Pueppke, 1992), luciferase (Cebolla et al., 1993; Pitkäjärvi et al., 2003), Green and DsRed fluorescent proteins (Duodu et al., 2008; Gage, 2002) and *CelB* (Sanchez-Canizares & Palacios, 2013; Sessitsch et al., 1996). Of these markers, *lacZ* has been used the longest to tag bacteria, but a high background activity in rhizobia restricts the use of this gene in rhizobial competition studies (Sessitsch et al., 1998). The main disadvantages of luciferase and fluorescent proteins are that they require sophisticated amplification devices and fluorescent microscopes (Sessitsch et al., 1998).

Two of the most robust gene markers used in gram-negative bacteria are the *Escherichia coli gusA* and the *Pyrococcus furiosus celB*. *gusA* encodes a β-glucuronidase (GUS) that cleaves glucuronides such as the synthetic 5-Bromo-4-chloro-3-indoxyl-beta-D-glucuronide cyclohexylammonium salt (X-Glc) into a colourless product that is further oxidised into an insoluble coloured product (Jefferson et al., 1986). As GUS is not expressed in rhizobia or by their legume hosts, *gusA* markers are ideal for rhizobia nodule occupancy studies (Sessitsch et al., 1996; Wilson et al., 1995). GUS assays are simple to conduct and sensitive enough to allow nodules occupied by *gusA* marked strains to be easily identifiable (Reeve et al., 1999; Wilson et al., 1995). On the other hand, *celB* encodes a thermostable and thermoactive β-glucosidase with a high β-

galactosidase activity (Voorhost et al., 1995). As rhizobia and their hosts contain heat-labile β-galactosidases, a heat inactivation step is used to eliminate endogenously expressed β-galactosidases before assay with β-galactosides (Sessitsch et al., 1996). The use of the two markers allows the simultaneous detection of nodule occupants on a root (Sanchez-Canizares & Palacios, 2013; Sessitsch et al., 1996).

Delivery of marker genes into rhizobia

Marker genes are introduced into rhizobia through different approaches. For example, *gusA* genes have been introduced into rhizobia through site-directed PCR methods (Sanchez-Canizares & Palacios, 2013) or with the aid of transposons (Reeve et al., 1999; Wilson et al., 1995) and plasmids (Shamseldin, 2007; Weilbo et al., 2010). In general, delivery methods should (a) be easy to use, (b) enable rapid delivery of genes, (c) be amenable to use with many strains, (d) lead to stable marker genes even in the absence of selection, (e) not result in a phenotype either by interrupting essential genes or by excessive metabolic burden, (f) be non-specific and allow marking of relatively uncharacterized rhizobial strains, and (g) permit the monitoring of rhizobia in the free-living state through suitable selection mechanisms.

Site-directed methods can deliver marker genes into specific targets e.g. by fusion PCR and double recombination (Sanchez-Canizares & Palacios, 2013). However, these are significantly laborious and therefore not amenable to the marking of many strains, thereby limiting utility for marking a substantial number of strains. Additionally, the method needs considerable knowledge on the genome of strain to be marked to allow for decisions on suitable insertions sites. Genome information of strains to be included in competition assays in not always available.

The current mode of choice for introducing marker genes into bacteria is through mini-transposons, and a range of these is now available for marking of rhizobia (Reeve et al., 1999; Wilson et al., 1995) (Figure 1.5). Although mini-transposons are widely used to introduce marker genes into rhizobia (Sessitsch et al., 1996), they suffer a few drawbacks. The major one being that they result in random integrations of introduced genes into the target genomes, potentially disrupting essential genes. Consequently, the phenotypes of the mutants have to be checked (De Lorenzo et al., 1998; Wilson et al., 1995) and preferably, the location of the transposon traced through PCR and sequencing techniques (Chun et al., 1997). The screening of mutants for mini-

transposon insertion sites is a laborious process and is a fundamental limitation to the use of transposons to mark strains. However, for the marking of a few strains, the delivery of marker genes through mini-transposons is a suitable approach (Reeve et al., 1999; Sessitsch et al., 1998; Wilson et al., 1995).

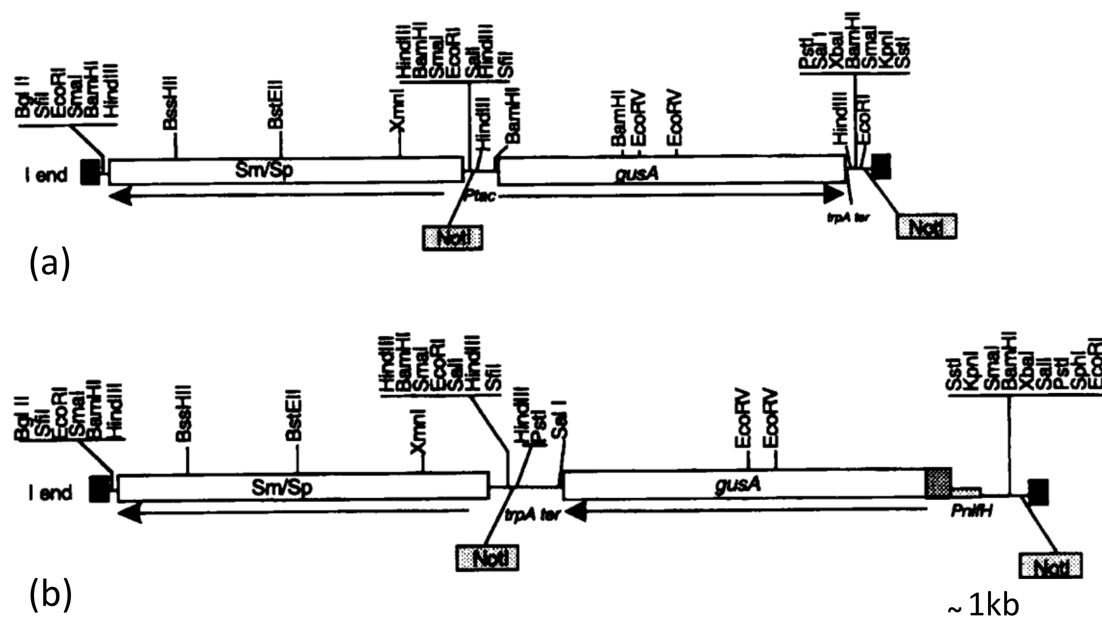


Figure 1.5: Restriction maps of GUS mini-transposons (a) mTn*5*SS*gusA*11, with *gusA* gene, expressed constitutively from a *tac* promoter (b) mTn*5*SS*gusA*31, with *gusA* from a symbiotically active *nifH* promoter. Unmodified from Wilson et al. (1995)

Plasmids are an alternative mode of introducing marker genes into rhizobia. While they are easy to conjugate into bacteria, their main disadvantages include instability in the absence of selection (Corich et al., 2001; De Gelder et al., 2007; Duodu et al., 2008; Gage et al., 1996) and the introduction of an extra metabolic load in the transconjugants (Silva et al., 2011).

A stable, broad-host-range plasmid used with rhizobia is pJP2 (Figure 1.6) (Prell et al., 2002). pJP2 is maintained stably in *R. leguminosarum* bacteroids (Karunakaran et al., 2005; Prell et al., 2002) while pTR102, a plasmid with the same broad host range features and stability region is also stably maintained in *Sinorhizobium meliloti* strains re-

isolated from alfalfa nodules (Weinstein et al., 1992). The introduction of these plasmids into rhizobia has been reported to result in no detrimental phenotype in the transconjugants (Prell et al., 2002) in part due to a relatively low copy number and a relatively weakly expressed *tac* promoter (Giacomini et al., 1994). pJP2 is a mini-RK2 derivative that has important RK2 elements that confer broad host range and stability. These are *oriV, trfA* (the replication initiation protein that allows replication in various hosts) and the *par* region (Jain & Srivastava, 2013; Yano et al., 2012).

pJP2 uses two stability mechanisms. One mechanism avoids plasmid loss while the other eliminates the survival of plasmid-free cells. The *parCBA* operon encodes an active partition system that distributes plasmids (spatially) within the cell to ensure daughter cells receive a plasmid upon division (Easter et al., 1998; Sobecky et al., 1996). A second mechanism involves post-segregational killing relating to a toxin and an antitoxin (TA) system. In this particular TA (type II) system, *parD* codes for an antitoxin that neutralises a toxin from *parE* (Goeders & Van Melderen, 2014). A daughter cell that does not receive a plasmid dies, as no antitoxin is available to bind to and neutralise the long lasting toxin (Jiang et al., 2002).

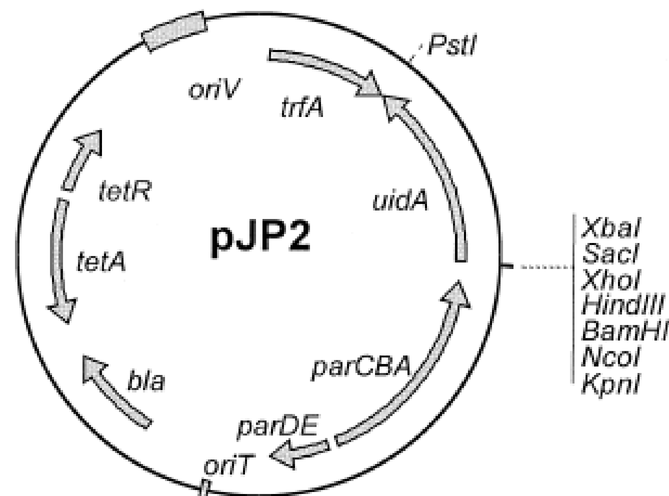


Figure 1.6: Physical and genetic map of the pJP2. The promoter-probe vector encodes genes for tetracycline (*tetA*/*tet*R) and ampicillin (*bla*) resistance. In addition to these, RK2 genes *trfA*, *oriV* and *oriT* are incorporated as well as *parCBA* and *parDE* operons for plasmid stability (Prell et al., 2002).

Although RK2 derivatives such as pJP2 carry an origin of transfer (*oriT*), they are non-conjugative as they lack *tra* genes (Bates et al., 1998) and therefore transconjugants can be grown together with other strains without the risk of transferring the marker gene. With the success of mini-RK2 derivatives in marking *R. leguminosarum* and *S. meliloti* (Karunakaran et al., 2005; Prell et al., 2002; Weinstein et al., 1992), a related mini-RK2 derivative would be a useful tool for introducing marker genes rapidly into rhizobia in competition studies, although the stability of such a plasmid in an uncharacterised, unknown collection of rhizobia would need to be tested.

1.5 *P. vulgaris* in Kenya

P. vulgaris has been cultivated in Kenya for 400-500 years, following its introduction by Portuguese traders at the East African coast (Greenway, 1944). The average yield of *P. vulgaris* in Kenya is low, remaining below 659 kg ha^{-1} between 2004 and 2014 (Figure 1.7). For comparison, the average yield in Brazil in 2014 was 1034 kg ha^{-1} (FAOSTAT). The major causes of poor yields in Kenya include low yielding cultivars, diseases, drought and low soil fertility (Katungi et al., 2009).

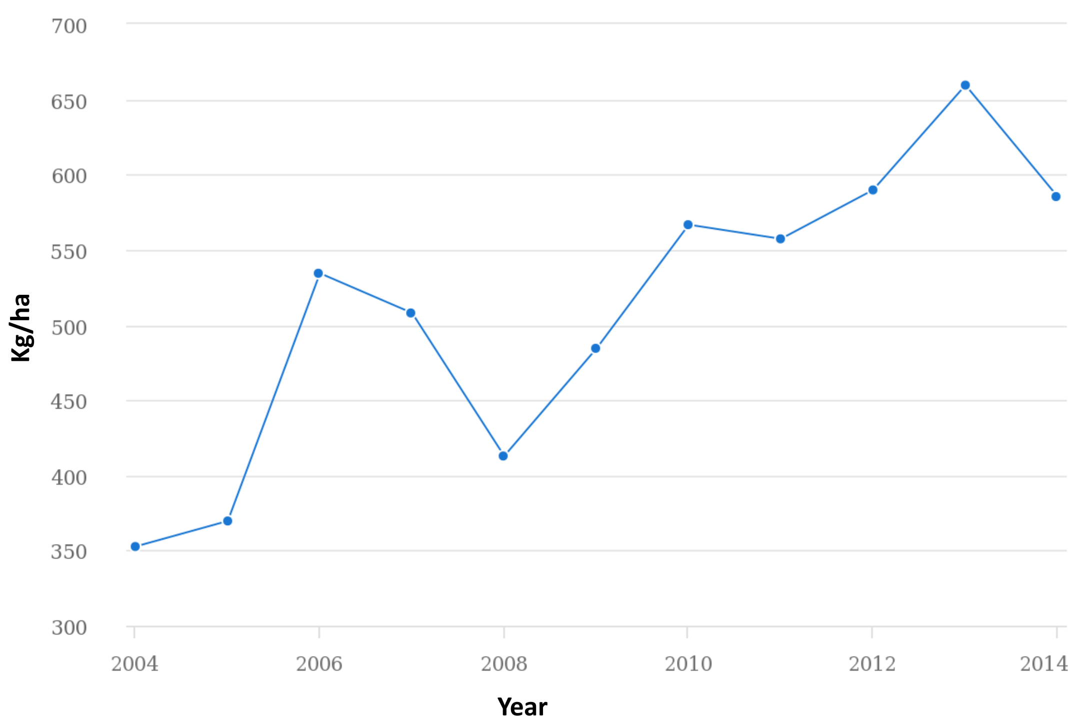


Figure 1.7: The average yield of *P. vulgaris* in Kenya in the years 2004-2014 (FAOSTAT)

Because low soil N is a major limiting factor to the growth of this crop in Kenya, there is potential to increase yields of *P. vulgaris* through N_2 fixation as has been achieved in

Brazil, where yield increases of up to 900 kg ha^{-1} were obtained through inoculation with an effective and ecologically adapted strain (Hungria et al., 2000). In Kenya, *P. vulgaris* is inoculated mainly with *R. tropici* CIAT 899 (Bala et al., 2011), and occasionally with *R. etli* USDA 2667 (Koinange, 2015). CIAT 899, the type strain for *Rhizobium tropici* was isolated from *P. vulgaris* in Colombia (Martinez-Romero et al., 1991) and is effective on *P. vulgaris* (Hungria et al., 2000; Hungria et al., 2003). The strain is tolerant to high temperatures and acidity (Hungria et al., 2000). Under acidic pH, CIAT 899 is competitive for nodulation of *P. vulgaris* (Hungria et al., 2003; Shamseldin, 2007; Streit et al., 1992) although more competitive strains have been reported (Anyango et al., 1998; Hungria et al., 2003). In near-neutral pH, CIAT 899 has been reported to be poorly competitive (Anyango et al., 1998; Streit et al., 1992). The strain is used as inoculant strain in Brazil (Hungria et al., 2000) and several countries in SSA (Bala et al., 2011).

The inoculation of *P. vulgaris* in Kenya is not widespread, and the success rate of inoculation is low, with most studies reporting no response to inoculation even in instances when the crop responded to the application of fertiliser N (Musandu & Ogendo, 2001). Reasons for inoculation failure are speculated to include a low adaptability of inoculant strains to environmental conditions, and the presence of large populations of indigenous rhizobia soils (Anyango et al., 1998; Musandu & Ogendo, 2001), although no study has assessed the competitiveness of inoculant strains against indigenous strains.

Populations of rhizobia nodulating *P. vulgaris* in Kenyan soils range upward to 10^4 cells g $g^{-1}$of soil (Anyango et al., 1995; Kawaka et al., 2014) and occasionally, some of these populations contain strains that are very efficient at fixing N_2 (Anyango et al., 1995; Karanja & Wood, 1988; Kawaka et al., 2014; Muthini et al., 2014; Mwenda et al., 2011). For example, Kawaka et al. (2014) isolated KSM005 that outperformed the commercial inoculant CIAT 899 in biomass production in controlled glasshouse trials. However, greater efforts are required to create a bigger pool of candidate inoculant strains for adaptability assessment in different agro-ecological zones in Kenya.

Despite the abundance of *P. vulgaris* rhizobia reported in Kenyan soils, and the potential of some of them to fix high amounts of N_2, very little else is known about them. For example, beyond morphological assessment (Karanja & Wood, 1988; Kawaka et al., 2014; Muthini et al., 2014), few researchers have attempted to classify rhizobia from Kenyan soils by molecular means (Anyango et al., 1995; Mwenda et al., 2011; Odee et

al., 2002). The most comprehensive of these was the study by Anyango et al. (1995) that by RFLP of genomic DNA and *nifH* copy number, found strains nodulating *P. vulgaris* in alkaline soils of Naivasha to be closely related to *R. etli* and *R. leguminosarum* bv. phaseoli and those nodulating *P. vulgaris* in acidic soils in Dakaini to be related to *R. tropici*. It is now 21 years since the study by Anyango et al. (1995) was published. Since 1995, many new rhizobial species that fix N_2 with *P. vulgaris* have been described (Table 1.1) and others such as *R. leucaenae* (Ribeiro et al., 2012), *R. freirei* (Dall'Agnol et al., 2013) and *R. paranaense* (Dall'Agnol et al., 2014) have since been separated from what was *R. tropici* in 1995. An updated view of the rhizobia nodulating *P. vulgaris* in Kenya is needed.

1.6 Aims of thesis

Several knowledge and resource gaps exist in the *P. vulgaris* symbiosis, in Kenya and in general. These gaps include a lack of effective, well adapted and competitive strains for use as inoculants in the diverse agroecological zones in Kenya, a lack of knowledge on the nature of rhizobia that presently nodulate *P. vulgaris* in Kenya and insufficient knowledge on the relative importance of factors currently believed to impact on the success of inoculating *P. vulgaris* in multi-strain environments.

To address the gaps outlined above, the specific aims of this thesis are:

1. To assess the genetic diversity and phylogeny of rhizobia nodulating *P. vulgaris* in select agro-ecological zones in Kenya
2. To determine the effectiveness, in N_2-fixation, of the strains on *P. vulgaris* and, additionally, explore any link between β-polyhdroxybutyrate accumulation in bacteroids and strain effectiveness
3. Develop and test a dual gene-marker system based on *gusA* and *celB* genes for use in rapid screening of rhizobia for competitiveness in nodulating *P. vulgaris*
4. Investigate the determinants of inoculation success in *P. vulgaris* in soils with high densities of resident rhizobia

CHAPTER 2
The genetic diversity of bacteria that nodulate *P. vulgaris* in Kenya

2.1 Introduction

In Kenya, low levels of soil N significantly limit the productivity of crops (Keino et al., 2016). With *P. vulgaris*, N_2 fixation has the potential to increase the current yields of below 700 kg ha^{-1} (Figure 1.7), but a deeper understanding of the *P. vulgaris*-rhizobia symbiosis in Kenyan soils is required. A critical step towards a greater understanding of any legume-rhizobia interaction is the genetic characterization of the rhizobia involved. Such characterization enhances knowledge on the genetic relationships among strains, the predominance of strains and the dynamics of exchange of genetic material.

P. vulgaris is cultivated in central and western Kenya, among other parts of the country (Katungi et al., 2009). Although inoculation of the crop is rare, nodulation with native rhizobia, whose population is estimated to be as high as 10^4 cells g^{-1} of soil, occurs (Anyango et al., 1995; Kawaka et al., 2014). Studies indicate these rhizobia are variably effective on *P. vulgaris* (Anyango et al., 1995; Kawaka et al., 2014; Muthini et al., 2014) but little else is known about them e.g. their diversity, taxonomy or their nodulation genes. Previously, rhizobia nodulating *P. vulgaris* at two sites in Kenya grouped with members of *R. etli*, *R. leguminosarum* bv. phaseoli and *R. tropici* by host range, *nifH* copy number and genomic DNA restriction fragment fingerprints (Anyango et al., 1995). In a second study, two strains from *P. vulgari*s were identified as *R. leguminosarum* using a 230 bp fragment of the 16S rRNA gene (Odee et al., 2002). To date, no other studies have genetically characterized the rhizobia that nodulate *P. vulgaris* in Kenyan soils.

To form a better view of the nature of rhizobia nodulating *P. vulgaris* in Kenyan soils, I examined a collection of 197 rhizobial strains by genomic DNA fingerprinting; restriction fragment length polymorphism (RFLP) of PCR amplified of 16S rRNA (PCR-RFLP); and by the sequencing of 16S rRNA, *recA* and *nodC* genes.

2.2 Materials and methods

2.2.1 Nodule collection, strain isolation and storage

Nodules were sampled from *P. vulgaris* growing in farms with no known history of rhizobial inoculation in Nairobi, Kiambu, Meru and Siaya. Nodules were stored in silica gel in airtight vials before isolation of bacteria. Isolation was as described by Hungria et al. (2016). Briefly, desiccated nodules were re-hydrated in de-ionized water for 3 h before surface sterilization in 70% (v/v) ethanol for 1 min followed by immersion in 4% (v/v) sodium hypochlorite for 2 min. Nodules were then rinsed six times in sterile de-ionized water and a nodule aseptically crushed in a drop of sterile de-ionized water before streaking of the 'squashate' on tryptone yeast (TY) media [0.5% tryptone (w/v), 0.3% yeast extract (w/v), 0.087% $CaCl_2.2H_2O$ (w/v), 0.75% agar (w/v), (pH 7.0)]. The streaked plates were incubated at 28ºC for 3 to 10 d, colonies sub-cultured onto fresh TY plates, and pure cultures stored in 15% (v/v) glycerol at -80ºC.

In addition to these isolates, 12 strains from the N2Africa Project's rhizobial collection were included in the study (Table 2.1). All but NAK 8, NAK 104, NAK 111, and NAK 120 had been isolated from *P. vulgaris* (Table 2.1).

Table 2.1: Rhizobial strains included in the study from N2Africa rhizobia collection at the Nairobi MIRCEN laboratory

Strain	Geographic origin*	Host of isolation
NAK 8	Embu (CK)	*Macroptilium atropurpureum***
NAK 73A	Nairobi (CK)	*Phaseolus vulgaris*
NAK 73B	Nairobi (CK)	*Phaseolus vulgaris*
NAK 75	Nairobi (CK)	*Phaseolus vulgaris*
NAK 103	Busia (WK)	*Phaseolus vulgaris*
NAK 104	Busia (WK)	*Phaseolus vulgaris*
NAK 105	Busia (WK)	*Glycine max*
NAK 108	Busia (WK)	*Phaseolus vulgaris*
NAK 111	Busia (WK)	*Vigna unguiculata*
NAK 120	Busia (WK)	*Albizia sp.*
NAK 156	Kwale (CST)	*Phaseolus vulgaris*
NAK 157	Kwale (CST)	*Phaseolus vulgaris*

* CK-central Kenya, WK-western Kenya, CST-coastal Kenya
** trap host

In all, the study obtained nodule samples or strains from the central highlands, western and coastal Kenya (Figure 2.1), regions that differ ecologically in elevation, rainfall, and soil types (Table 2.2).

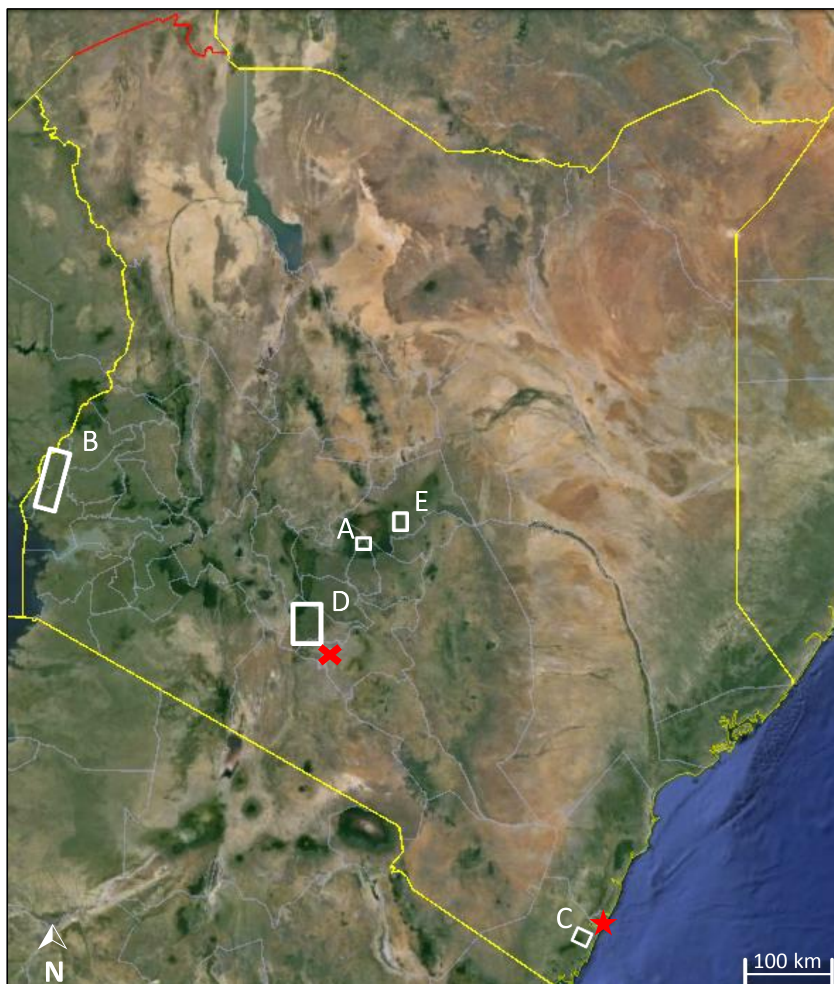


Figure 2.1: Map of Kenya showing geographical origin of nodule samples or strains: **A** (Embu), **B** (Busia and Siaya), **C** (Kwale), **D** (Kiambu and Nairobi) and **E** (Meru). For more details on these areas, see Table 2.2. The main cities in Kenya, Nairobi (✖) and Mombasa (★) are shown. Map generated with Google Earth.

Table 2.2: A brief description of the geographic origins of nodules or strains, indicating administrative areas, altitudes, annual rainfall, and soil types

Zone[a]	Administrative area (county)[b]	Approximate altitudes, annual rain and soil types[c]
A	Embu (CK)	2000 m.a.s.l, 1500 mm, humic nitisols
B	Busia/Siaya (WK)	1200-1500 m.a.s.l, 1200-2000 mm, acrisols and ferralsols
C	Kwale (CST)	0-100 m.a.s.l, <500 mm, low fertility arenosols
D	Kiambu/Nairobi (CK)	1600-2000 m.a.s.l, 1200-1800 mm, humic nitisols
E	Meru (CK)	1700 m.a.s.l, 1500-2000 mm, cambisols

[a]Zones corresponding to those shown in Figure 2.1
[b]CK-central Kenya, WK-western Kenya, CST-coast. Table 2.3 gives more specific locations of sampling points
[c]Generalized descriptions of sampling points

2.2.2 Strain authentication

Isolates were first assessed for growth rates then authenticated. To determine growth rates, isolates were dilution-streaked onto TY plates, incubated at 28ºC and the time taken for colonies to appear recorded.

Authentication experiments were conducted in a glasshouse under natural light during the months of October to March in Perth, Western Australia. These are months of bright sun with 12 to 14-h days and glasshouse air temperatures ranging between 22°C and 30°C. General procedures were as described by Yates et al. (2016a). Briefly, 3.5 L plastic pots were filled with wet medium-grade vermiculite and steam-sterilised for 3 h after which pots were covered with cling wrap and kept in a shaded glasshouse until sowing. *P. vulgaris* cv. KK08 seeds were selected for uniformity in size and weight, surface-sterilized by immersion in 70% (v/v) ethanol for 1 min and in 4% (v/v) sodium hypochlorite for 4 min before rinsing in six changes of sterile de-ionized water. The seeds were allowed to imbibe water for 2 h before pre-germination on water agar [0.75% (w/v) agar] in a 28°C incubator for 48 h.

Strains to be authenticated were inoculated, from TY agar plates, into 10 mL TY broths and incubated at 28°C on a gyratory shaker set at 220 rpm and grown until early stationary phase. The cultures were centrifuged at 10,000 × *g* for 1 min, supernatant removed and cell pellets re-suspended in 10 mL sterile de-ionized water. The optical density (OD) of the resulting bacterial suspensions was measured at 600 nm and OD adjusted to 0.5 using sterile de-ionized water to obtain approximately 5×10^8 cells mL^{-1}.

At the time of planting, pre-germinated seeds were aseptically placed into shallow holes in the vermiculite, and 1 mL of the freshly prepared cell suspensions applied onto each seed. Three seeds were sown per pot and treatments had two pot replicates. Non-inoculated control treatments received 1 mL of sterile de-ionized water at sowing. A capped watering tube was inserted into each pot and the pot surface covered with a thin layer of sterile polyethylene beads to prevent airborne contamination. Each pot was watered with 100 mL of 2 × B&D nutrient solution ($CaCl_2$, 2 mM; KH_2PO_4, 1 mM; ferric citrate, 20 μM; $MgSO_4$, 0.5 mM; K_2SO_4, 0.5 mM; $MnSO_4$, 2 μM; H_3BO_3, 4 μM; $ZnSO_4$, 1 μM; $CuSO_4$, 0.4 μM; $CoSO_4$, 0.2 μM; Na_2MoO_4, 0.1 μM) (Broughton & Dilworth, 1971) supplemented with 10 mM NH_4NO_3 as starter N. Plants were thinned to two plants per pot 7 d after sowing. 150 mL of N-free B&D nutrient solution was

added to pots twice per week, and pots watered with sterile distilled water as needed until harvest. Plants were assessed for root nodulation 28 d after inoculation.

2.2.3 Genetic characterization

2.2.3.1 Genomic DNA preparation

Genomic DNA was extracted according to Baele et al. (2000). Briefly, strains were grown on TY agar plates and a bacterial colony suspended in 20 μL of lysis buffer (0.25% (w/v) sodium dodecyl sulfate, 0.05 M NaOH) before incubation at 95ºC for 5 min. The lysate was centrifuged for 1 min at 16,000 × *g* followed by addition of 180 μL sterile de-ionized water. The resulting suspension was centrifuged at 16,000 × *g* for 5 min to pellet cell debris, and the supernatant containing genomic DNA stored at -20ºC.

2.2.3.2 Genomic DNA fingerprinting by PCR with *nif*-directed RP01 primer

PCR was carried out using the *nif*-directed RP01 primer '5-AATTTTCAAGC GTCGTGCCA-3' (Richardson et al., 1995). As a control, two independent reactions of NAK 73 were set up to confirm the reproducibility of the method. A single reaction was carried out for all other strains. PCR reactions of 20 μL were performed with 4 μL of 5 × PCR polymerization buffer from Fisher Biotec [composition of 1 × buffer: 67 mM Tris-HCl pH 8.8, 16.6 mM $[NH_4]_2SO_4$, 0.45% (v/v) Triton X-100, 0.2 mg/mL gelatin and 0.2 mM dNTP's], 3mM $MgCl_2$, 1.25 μM RP01 primer, 2.5 U Taq DNA polymerase (Fisher Biotec), 1 μL DNA and UltraPure PCR grade water (Fisher Biotec) to 20 μL. The thermal cycling conditions for the PCR were: 5 cycles at 94ºC for 30 s, 50ºC for 10 s and 72ºC for 90 s; 35 cycles of 94°C for 30 s, 55°C for 25 s and 72°C for 90 s; and a final extension at 72°C for 5 min.

The entire 20 μL PCR reactions were each mixed with 3.2 μL of 6 × loading dye (Thermo Fisher Scientific) then separated on 2% (w/v) agarose gel in 1 × TAE buffer (40 mM Tris, 20mM acetic acid, 1 mM EDTA) at 80 V for 180 min. Three lanes on each gel were loaded with 7 μL of 1 kb DNA ladder (Promega, catalogue # G5711) to act as size markers. Gels were pre-stained with SYBR® Safe DNA gel stain (1:10,000). PCR products were visualised using a UV transilluminator before the capture of a digital image using GEL-DOC 2000 (Bio-Rad). The resulting fingerprints were analysed in Bionumerics v5.1 (Applied Maths, Belgium). The similarity among digitized profiles was calculated using the Pearson correlation, and an average linkage (UPGMA) dendrogram

was derived from the patterns. Strains showing >80% similarity in their banding patterns were grouped together as 'RP01-PCR groups'.

2.2.3.3 Restriction Fragment Length Polymorphisms (RFLP)-PCR of 16S rRNA genes

The 16S rRNA genes were amplified from representative strains chosen from each of the RP01 groups, *R. tropici* CIAT 899, and *R. leguminosarum* 8002 using the primer pair 27F ('5-AGAGTTTGATCCTGGCTCAG-3') and 1492R ('5-ACGGCTACCTTGTTA CGACTT-3) (Lane, 1991). Each 25 μL PCR reaction contained: 12.5 μL of 2 × GoTaq® Green Master Mix [(pH 8.5), 400 μM dATP, 400 μM dGTP, 400 μM dCTP, 400 μM dTTP and 3 mM MgCl2], 0.4 μM 27F primer, 0.4 μM 1492R primer, 2 μL genomic DNA and UltraPure PCR grade water (Fisher Biotec) to 25 μL. The thermal cycling conditions for the PCR were: initial denaturation at 94ºC for 120 s; 35 cycles of 94°C for 30 s, 55°C for 25 s and 72°C for 90 s; and a final extension at 72°C for 7 min. The PCR products were mixed with loading dye and separated on a pre-stained 1.5% (w/v) agarose gel in 1 × TAE buffer at 100 V for 60 min as described in Section 2.2.3.2. Bands of ~1500 bp were excised with a blade under a blue light transilluminator then purified with Promega's Wizard®SV Gel and PCR clean-up system (catalogue # A1330) following manufacturer's instructions.

Purified PCR products were digested separately with the restriction enzymes HaeIII, MspI, HhaI and HinfI (all from Promega) as recommended by the manufacturer. Briefly, digestion was carried out in 20 μL reactions that contained: 1 × restriction buffer (recommended and supplied with the restriction enzymes), 0.1 μg/μL acetylated BSA, 0.25 U/μL of an enzyme, and ~0.5 μg DNA. Digestion was carried out at 37°C for 4 h followed by separation in 2% (w/v) agarose at 80 V for 2 h, as already described. A digital image of the gel was captured using GEL-DOC 2000 (Bio-Rad) and banding patterns scored. A combination of RFLP patterns of the 16S rRNA fragments generated by the enzymes was used to group the isolates into 'PCR-RFLP groups'.

2.2.3.4 Amplification and sequencing of 16S rRNA gene

Near full-length 16S rRNA genes were amplified from representatives of RFLP groups and purified as described in Section 2.2.3.3. Sequencing PCR was then performed, separately, with the 27F and 1492R primers using the Bigdye® Terminator v3.1 Cycle Sequencing Kit (catalogue # 4337454) as per Applied Biosystems' instructions. Briefly, sequencing PCR was carried out in 10 μL reactions that contained: 1 μL of 2.5 ×

Bigdye® terminator ready reaction mix, 1.5 μL of 5 × Bigdye® sequencing buffer, 0.6 μM primer, ~90 ng DNA and UltraPure PCR grade water (Fisher Biotec) to 10 μL. The thermal cycling conditions for the PCR were: initial denaturation at 96ºC for 120 s; 25 cycles of 96°C for 10 s, 50°C for 5 s and 60°C for 240 s. Sequencing reactions were purified by ethanol, EDTA and sodium acetate precipitation following protocol by Applied Biosystems. Sequence reads were obtained from an ABI model 377A automated sequencer (Applied Biosystems), manually edited and contigs assembled in Geneious® software (Biomatters Ltd, NZ). Sequences were then used to search for bacterial type strains with the highest 16S rRNA gene similarity values in EzTaxon-e (Kim et al., 2012) and these sequences imported into MEGA6. A phylogenetic tree was constructed using the Maximum Likelihood method (with best fit models) bootstrapped with 500 replicates (Tamura et al., 2011).

2.2.3.5 Amplification and sequencing of *recA* gene

Partial *recA* genes (~550 bp) were amplified from the strains analysed by 16S rRNA gene above, using recA6F (5-CGKCTSGTAGAGGAYAAATCGGTGGA-3) and recA555R (5-CGRATCTGGTTGATG AAGATCACCAT-3) primers (Martens et al., 2007). PCR reactions of 25 μL were performed with 12.5 μL of 2 × GoTaq® Green Master Mix, 0.4 μM recA6F primer, 0.4 μM recA555R primer, 2 μL genomic DNA and UltraPure PCR grade water to 25 μL. The thermal cycling conditions for the PCR were: initial denaturation at 95ºC for 120 s; 32 cycles of 94°C for 45 s, 60°C for 60 s and 72°C for 90 s; and a final extension at 72°C for 5 min. Separation of PCR products on 1% (w/v) agarose gel and purification were as described in Section 2.2.3.3. Purified amplicons were sequenced by Sanger technique using recA6F primer through the Australian Genome Research Facility (AGRF) followed by contig assembly on Geneious®. *recA* sequences of test strains and those of highly similar type strains from GenBank were imported into MEGA6, aligned by Muscle and phylogeny reconstructed by the Maximum Likelihood method (with best fit models), bootstrapped with 500 replicates.

2.2.3.6 Amplification and sequencing of *nodC* gene

Partial *nodC* genes (~620 bp) were amplified using nodCfor540F (5-TGATYGAYATG GARTAYTGGCT-3) and nodCrev1160R (5-CAAYAGCGAYTGGYTRTC-3) primers (Sarita et al., 2005). PCR reaction 25 μL were performed with 12.5 μL of 2 × GoTaq® Green Master Mix, 0.4 μM nodCfor540F, 0.4 μM nodCrev1160R primer, 2 μL genomic

DNA and UltraPure PCR grade water to 25 μL. Thermal cycling conditions for the PCR were: initial denaturation at 95°C for 120 s; 3 cycles of 95°C for 60 s, 50°C for 135 s and 72°C for 60 s; 30 cycles of 94°C for 35 s, 50°C for 75 s and 72°C for 75 s; and a final extension at 72°C for 7 min. Separation of PCR products on 1% (w/v) agarose gel and purification were as described in Section 2.2.3.3. Purified amplicons were sequenced by Sanger technique using nodCfor540F primer through the Australian Genome Research Facility (AGRF) followed by contig assembly on Geneious®. The *nodC* sequences of test strains and those of highly similar type strains from GenBank were imported into MEGA6, aligned by Muscle and phylogeny reconstructed by the Maximum Likelihood method (with best fit models), bootstrapped with 500 replicates.

2.3 Results

2.3.1 Authenticated strains

A total of 185 isolates were obtained from the sampled nodules for a total of 197 study strains (with the addition of 12 strains from the N2Africa rhizobial collection) (Table 2.3). The origins of these isolates were; one strain from Embu (CK), 17 from Nairobi (CK), 25 from Meru (CK), 150 from Kiambu (CK), 12 from Busia/Siaya (WK) and two from Kwale (CK) (Table 2.3). On TY agar, an overwhelming majority of the strains (195 of the 197) were fast growers (taking 1-4 d for colonies to appear). Only 2 strains, NAK 105 and NAK 111, were slow growers (5-10 d for colonies to appear).

Pure cultures of all isolates induced nodules on *P. vulgaris* and were therefore authenticated as rhizobia. The uninoculated controls did form any nodules, an observation that further supported the authentication of the 197 study strains as rhizobia.

Table 2.3: A list of rhizobial strains used in this study. Their geographical origins, original hosts, growth rate phenotypes and subsequent characterizations

Isolate code*	Geographical origin in Kenya‡	Climate†	Host legume	Growth rate#	Authenticated$	RPO1-PCR	RFLP-PCR	16 S sequencing	*recA* sequencing	*nodC* sequencing
NAK 8*	Embu (CK)	TH	*M. atropurpureum*[¥]	F	✓	✓	–	–	–	–
NAK 73A*	Runda, Nairobi (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 73B*	Runda, Nairobi (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 75*	Runda, Nairobi (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 103*	Butula, Busia (WK)	TM	*P. vulgaris*	F	✓	✓	✓	✓	✓	✓
NAK 104*	Butula, Busia (WK)	TM	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 105*	Butula, Busia (WK)	TM	*G. max*	S	✓	✓	✓	–	–	–
NAK 108*	Butula, Busia (WK)	TM	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 111*	Butula, Busia (WK)	TM	*V. unguiculata*	S	✓	✓	✓	–	–	–
NAK 120*	Busia (WK)	TM	*Albizia sp.*	F	✓	✓	✓	✓	✓	✓
NAK 156*	Kwale (CST)	TCP	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 157*	Kwale (CST)	TCP	*P. vulgaris*	F	✓	✓	✓	✓	✓	–
NAK 209	Loresho, Nairobi (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 210	Loresho, Nairobi (CK)	STH	*P. vulgaris*	F	✓	✓	✓	✓	✓	✓
NAK 211	Loresho, Nairobi (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 212	Loresho, Nairobi (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 213	Loresho, Nairobi (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 214	Loresho, Nairobi (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 215	Loresho, Nairobi (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 216	Loresho, Nairobi (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 217	Loresho, Nairobi (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 218	Loresho, Nairobi (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 219	Loresho, Nairobi (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 220	Loresho, Nairobi (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 221	Loresho, Nairobi (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 222	Loresho, Nairobi (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 223	Kabete, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 224	Kabete, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 225	Kabete, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 226	Kabete, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 227	Kabete, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 228	Kabete, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 229	Kabete, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 230	Kabete, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 231	Kabete, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 232	Kabete, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 233	Kabete, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 234	Kabete, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 235	Kabete, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 236	Kabete, Kiambu (CK)	STH	*P. vulgaris*	F	✓	–	–	–	–	–
NAK 237	Kabete, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 238	Kabete, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 239	Kabete, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	✓	✓	✓
NAK 240	Kabete, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 241	Kabete, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–

Table 2.3 cont'd

NAK 242	Kabete, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 243	Kabete, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 244	Kabete, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 245	Wangige, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	✓	✓	✓
NAK 246	Wangige, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 247	Wangige, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 248	Wangige, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 249	Wangige, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 250	Wangige, Kiambu (CK)	STH	*P. vulgaris*	F	✓	–	–	–	–	–
NAK 251	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 252	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 253	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 254	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 255	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 256	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 257	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 258	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 259	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 260	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 261	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 262	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 263	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 264	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 265	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 266	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	✓	✓	✓
NAK 267	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 268	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 269	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 270	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 271	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 272	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 273	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 274	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 275	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 276	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 277	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 278	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 279	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 280	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 281	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 282	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 283	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 284	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓		–	–
NAK 285	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 286	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 287	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	✓	✓	✓
NAK 288	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 289	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	–	–	–	–	–
NAK 290	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 291	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 292	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	–	–	–	–	–
NAK 293	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 294	Ndenderu, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	✓	✓	✓

Table 2.3 cont'd

NAK 295	Muchatha, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 296	Muchatha, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 297	Muchatha, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 298	Muchatha, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 299	Muchatha, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 300	Muchatha, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 301	Muchatha, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 302	Muchatha, Kiambu (CK)	STH	*P. vulgaris*	F	✓	–	–	–	–	–
NAK 303	Muchatha, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 304	Muchatha, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 305	Muchatha, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 306	Muchatha, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 307	Muchatha, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 308	Muchatha, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 309	Muchatha, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 310	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 311	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 312	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	✓	✓	✓
NAK 313	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 314	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 315	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 316	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 317	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 318	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 319	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 320	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 321	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 322	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 323	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	–	–	–	–	–
NAK 324	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 325	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 326	Tigoni, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 327	Tigoni, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 328	Tigoni, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 329	Tigoni, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 330	Tigoni, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 331	Tigoni, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 332	Tigoni, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 333	Tigoni, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 334	Tigoni, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	✓	✓	✓
NAK 335	Tigoni, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 336	Tigoni, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 337	Tigoni, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 338	Tigoni, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 339	Tigoni, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 340	Tigoni, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 341	Tigoni, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 342	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 343	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 344	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 345	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 346	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 347	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–

Table 2.3 cont'd

NAK 348	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	–	–	–	–	–
NAK 349	Limuru, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	✓	✓	✓
NAK 350	Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 351	Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	✓	✓	–
NAK 352	Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 353	Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 354	Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	✓	–	–
NAK 355	Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 356	Ruaka, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 357	Ruaka, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 358	Ruaka, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	✓	✓	✓
NAK 359	Ruaka, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 360	Ruaka, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 361	Ruaka, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 362	Ruaka, Kiambu (CK)	STH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 363	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 364	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 365	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 366	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 367	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 368	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 369	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 370	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 371	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	–	–	–	–	–
NAK 372	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 373	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 374	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 375	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 376	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 377	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 378	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 379	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 380	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 381	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 382	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	✓	✓	✓	✓
NAK 383	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 384	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	✓	✓	✓	–
NAK 385	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	✓	–	–	–
NAK 386	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	–	–	–	–
NAK 387	Kinoru,Meru (CK)	TH	*P. vulgaris*	F	✓	✓	✓	✓	✓	✓
NAK 403	Sidada, Siaya (WK)	TM	*P. vulgaris*	F	✓	–	–	–	–	–
NAK 405	Sidada, Siaya (WK)	TM	*P. vulgaris*	F	✓	–	–	–	–	–
NAK 407	Sidada, Siaya (WK)	TM	*P. vulgaris*	F	✓	–	–	–	–	–
NAK 440	Sidada, Siaya (WK)	TM	*P. vulgaris*	F	✓	–	–	–	–	–
NAK 441	Sidada, Siaya (WK)	TM	*P. vulgaris*	F	✓	–	–	–	–	–
NAK 458	Sidada, Siaya (WK)	TM	*P. vulgaris*	F	✓	–	–	✓	✓	✓

*Isolates from the N2Africa rhizobia collection; ‡locality, administrative county and generalised location within Kenya (CK: central Kenya, WK-western Kenya, and CST-coastal Kenya); †STH-subtropical highland, TH-tropical highland, TCP-tropical coastal plain, TM-tropical midland; ¥Trap host; #Growth rate on YMA/TY media: F (1-4 d to appear) S (5-10 d to appear); $Authenticated as rhizobia on *P. vulgaris*; ✓ done, – not determined

2.3.2 Diversity of strains based on RP01-PCR

One hundred and eighty three (183) of the 197 isolates were successfully typed by RP01-PCR. PCR reactions with 14 isolates did not result in products that could be analyzed further. The successful PCR amplifications generated 5-15 bands following electrophoresis (Figure 2.2), resulting in banding patterns that differed by strain. Similarities in banding patterns ranged from as low as 15% to as high as 98%. The duplicate reactions of NAK 73 had banding patterns with a similarity of 92%. Based on the similarity of banding patterns of the NAK 73 duplicates, and previously observed reproducibility of RP01-PCR and DNA fingerprints (Gevers et al., 2001; Richardson et al., 1995), strains showing greater than 80% banding pattern similarity were grouped together, resulting in 88 RP01-PCR groups (Figure 2.2). The high number of RP01-PCR groups was indicative of a large genetic diversity among the study strains.

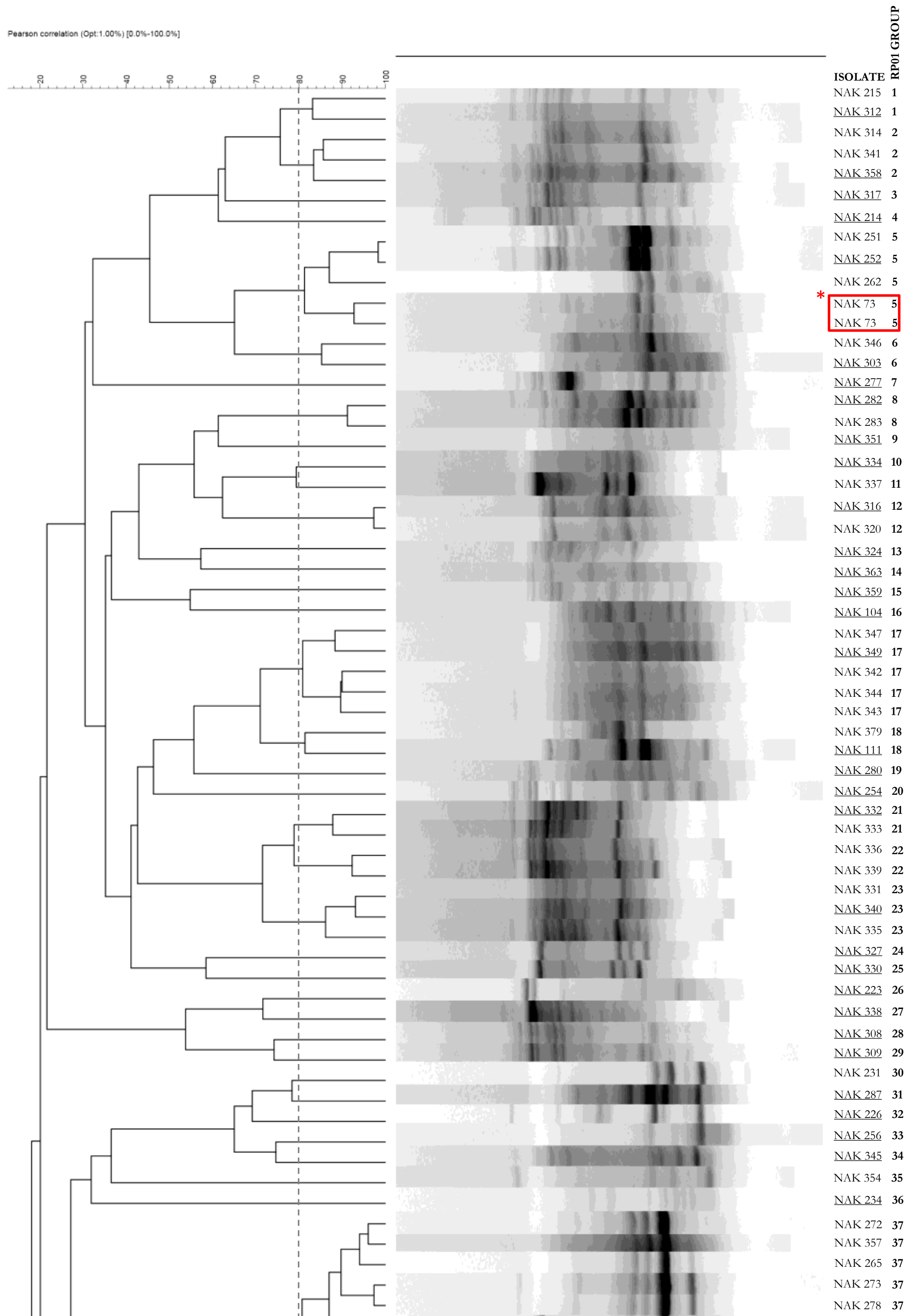
Additional analysis revealed that the similarity of strains by RP01-PCR was sometimes linked to their geographic origin. For example, 7 of 16 strains from Tigoni (Groups 21-23) had RP01-PCR banding patterns with greater than 70% similarity, as had 10 of 25 strains from Meru (Groups 69-70). Contrastingly, some RP01-PCR groups (≥70 similarity) contained strains from diverse locations. For example, in RP01-PCR group 50, NAK 103 isolated from Busia in western Kenya, had a banding profile with an over 90% similarity to nine strains from Nairobi and Kiambu in central Kenya. This finding indicated that in some instances, strains similar by RP01-PCR were localized to sampling regions, while in other instances similar strains were found in geographically disparate regions.

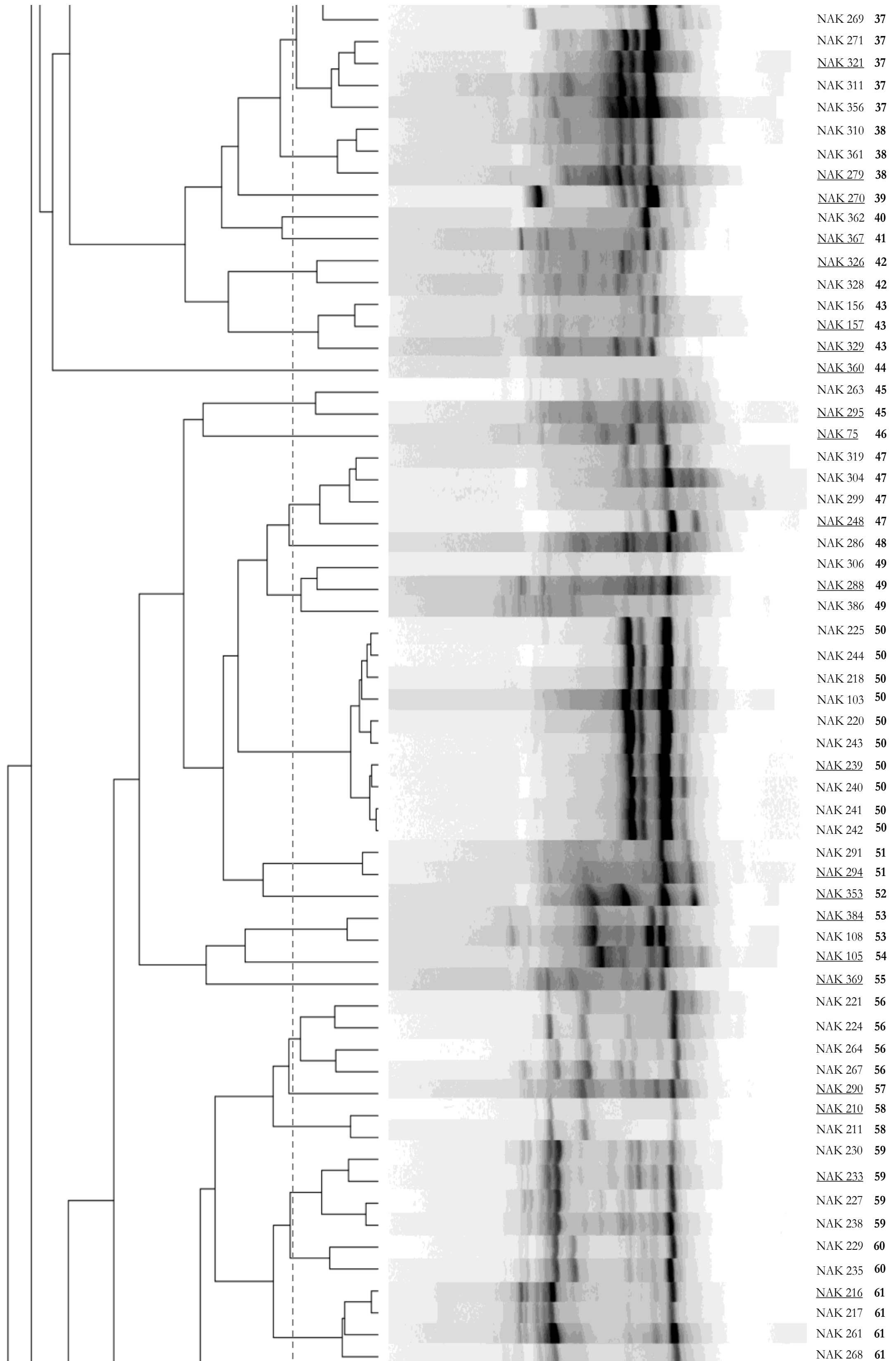
2.3.3 Diversity of strains based on RFLP of PCR-amplified 16S rRNA

At least 1,450 bp of 16S rRNA gene sequence was amplified for 85 strains representative of almost all RP01-PCR groups (Section 2.3.2). Digestion of the 85 amplicons separately with MspI, HaeIII, HinfI and HhaI and gel electrophoresis of the products resulted in 3-8 distinct banding patterns per restriction enzyme (Table 2.4). MspI had the greatest power to resolve differences in the 16S rRNA genes, giving eight different banding patterns (a-h). Digestion with HaeIII and HinfI resulted in 4 patterns (a-d) while HhaI gave 3 patterns (a-c). By combining the restriction patterns obtained with each of the four enzymes, the study strains generated nine PCR-RFLP groups designated 1-9 (Table 2.4).

A total of 55 strains, including the reference *R. leguminosarum* strain 8002, shared identical banding patterns for all four restriction enzymes and were grouped together (PCR-RFLP Group 1). Twenty strains belonged to PCR-RFLP Group 2 while four strains that included *R. tropici* CIAT 899 belonged to PCR-RFLP Group 3. Each of the remaining PCR-RFLP groups (4-9) had a single strain.

Following the PCR-RFLP analysis, 18 strains were chosen for sequencing of near full-length 16S rRNA genes. The choice of the strains was based on their representativeness of the PCR-RFLP groups, geographic origins and their perceived value in subsequent chapters on effectiveness and competition. The value was assessed using data from the visual assessment of the symbiotic performance of strains on *P. vulgaris* during authentication.


Pearson correlation (Opt:1.00%) [0.0%-100.0%]
ISOLATE
RP01 GROUP
NAK 215 1
NAK 312 1
NAK 314 2
NAK 341 2
NAK 358 2
NAK 317 3
NAK 214 4
NAK 251 5
NAK 252 5
NAK 262 5
NAK 73 5
NAK 73 5
NAK 346 6
NAK 303 6
NAK 277 7
NAK 282 8
NAK 283 8
NAK 351 9
NAK 334 10
NAK 337 11
NAK 316 12
NAK 320 12
NAK 324 13
NAK 363 14
NAK 359 15
NAK 104 16
NAK 347 17
NAK 349 17
NAK 342 17
NAK 344 17
NAK 343 17
NAK 379 18
NAK 111 18
NAK 280 19
NAK 254 20
NAK 332 21
NAK 333 21
NAK 336 22
NAK 339 22
NAK 331 23
NAK 340 23
NAK 335 23
NAK 327 24
NAK 330 25
NAK 223 26
NAK 338 27
NAK 308 28
NAK 309 29
NAK 231 30
NAK 287 31
NAK 226 32
NAK 256 33
NAK 345 34
NAK 354 35
NAK 234 36
NAK 272 37
NAK 357 37
NAK 265 37
NAK 273 37
NAK 278 37


NAK 269 37
NAK 271 37
NAK 321 37
NAK 311 37
NAK 356 37
NAK 310 38
NAK 361 38
NAK 279 38
NAK 270 39
NAK 362 40
NAK 367 41
NAK 326 42
NAK 328 42
NAK 156 43
NAK 157 43
NAK 329 43
NAK 360 44
NAK 263 45
NAK 295 45
NAK 75 46
NAK 319 47
NAK 304 47
NAK 299 47
NAK 248 47
NAK 286 48
NAK 306 49
NAK 288 49
NAK 386 49
NAK 225 50
NAK 244 50
NAK 218 50
NAK 103 50
NAK 220 50
NAK 243 50
NAK 239 50
NAK 240 50
NAK 241 50
NAK 242 50
NAK 291 51
NAK 294 51
NAK 353 52
NAK 384 53
NAK 108 53
NAK 105 54
NAK 369 55
NAK 221 56
NAK 224 56
NAK 264 56
NAK 267 56
NAK 290 57
NAK 210 58
NAK 211 58
NAK 230 59
NAK 233 59
NAK 227 59
NAK 238 59
NAK 229 60
NAK 235 60
NAK 216 61
NAK 217 61
NAK 261 61
NAK 268 61

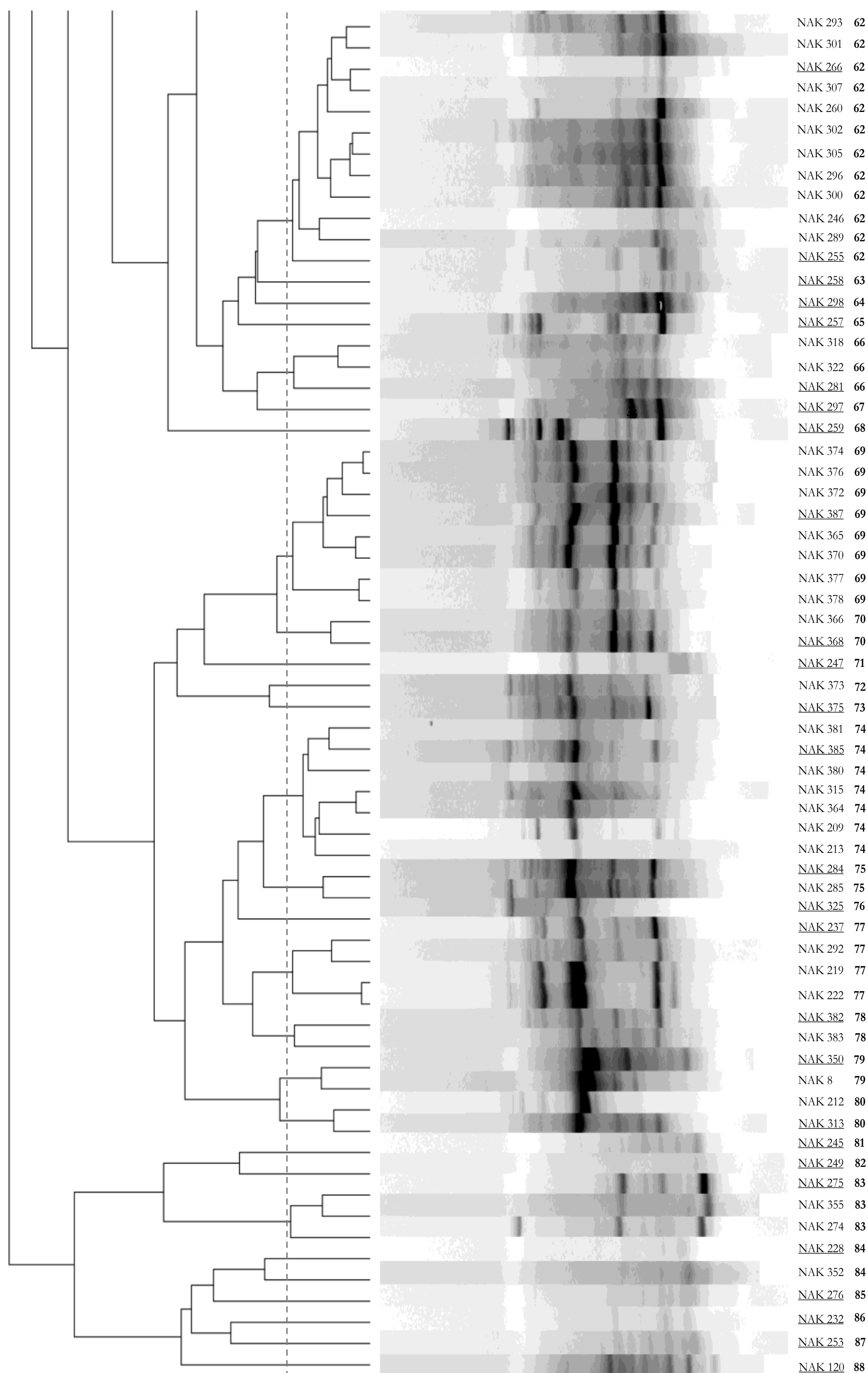


Figure 2.2: Dendrogram generated after cluster analysis of the RP01-PCR fingerprints of isolates. The dendrogram was constructed using the UPGMA method using arithmetic averages with correlation levels expressed as percentage values of the Pearson correlation coefficient. *Duplicates of NAK 73 as a control.

Table 2.4: Groupings of rhizobia generated by RFLP analysis of 16S rRNA genes

Strain‡	RP01 group †	16S restriction pattern*				RFLP 16S rRNA group#
		MspI	Hha1	HinfI	HaeIII	
8002	nd	a	a	a	a	1
NAK 210	58	a	a	a	a	1
NAK 214	4	a	a	a	a	1
NAK 216	61	a	a	a	a	1
NAK 233	59	a	a	a	a	1
NAK 237	77	a	a	a	a	1
NAK 247	71	a	a	a	a	1
NAK 252	5	a	a	a	a	1
NAK 253	87	a	a	a	a	1
NAK 254	20	a	a	a	a	1
NAK 255	62	a	a	a	a	1
NAK 257	65	a	a	a	a	1
NAK 258	63	a	a	a	a	1
NAK 259	68	a	a	a	a	1
NAK 266	62	a	a	a	a	1
NAK 270	39	a	a	a	a	1
NAK 276	85	a	a	a	a	1
NAK 277	7	a	a	a	a	1
NAK 279	38	a	a	a	a	1
NAK 280	19	a	a	a	a	1
NAK 281	66	a	a	a	a	1
NAK 282	8	a	a	a	a	1
NAK 284	75	a	a	a	a	1
NAK 288	49	a	a	a	a	1
NAK 290	57	a	a	a	a	1
NAK 295	45	a	a	a	a	1
NAK 297	67	a	a	a	a	1
NAK 298	64	a	a	a	a	1
NAK 303	6	a	a	a	a	1
NAK 308	28	a	a	a	a	1
NAK 309	29	a	a	a	a	1
NAK 312	1	a	a	a	a	1
NAK 313	80	a	a	a	a	1
NAK 316	12	a	a	a	a	1
NAK 317	3	a	a	a	a	1
NAK 321	37	a	a	a	a	1
NAK 324	13	a	a	a	a	1
NAK 325	76	a	a	a	a	1
NAK 326	42	a	a	a	a	1
NAK 327	24	a	a	a	a	1
NAK 329	43	a	a	a	a	1
NAK 330	25	a	a	a	a	1
NAK 332	21	a	a	a	a	1
NAK 334	10	a	a	a	a	1
NAK 338	27	a	a	a	a	1
NAK 340	23	a	a	a	a	1
NAK 345	34	a	a	a	a	1
NAK 358	2	a	a	a	a	1
NAK 359	15	a	a	a	a	1
NAK 363	14	a	a	a	a	1
NAK 367	41	a	a	a	a	1
NAK 369	55	a	a	a	a	1
NAK 375	73	a	a	a	a	1
NAK 382	78	a	a	a	a	1
NAK 385	74	a	a	a	a	1
NAK 103	50	a	a	b	b	2
NAK 105	54	a	a	b	b	2
NAK 157	43	a	a	b	b	2
NAK 223	26	a	a	b	b	2
NAK 226	32	a	a	b	b	2
NAK 228	84	a	a	b	b	2
NAK 232	86	a	a	b	b	2
NAK 234	36	a	a	b	b	2
NAK 239	50	a	a	b	b	2
NAK 245	81	a	a	b	b	2
NAK 248	47	a	a	b	b	2
NAK 249	82	a	a	b	b	2
NAK 256	33	a	a	b	b	2
NAK 275	83	a	a	b	b	2
NAK 287	31	a	a	b	b	2
NAK 350	79	a	a	b	b	2
NAK 351	9	a	a	b	b	2
NAK 360	44	a	a	b	b	2
NAK 368	70	a	a	b	b	2
NAK 387	69	a	a	b	b	2
CIAT 899	nd	c	b	a	a	3
NAK 120	88	c	b	a	a	3
NAK 349	17	c	b	a	a	3
NAK 104	16	c	b	a	a	3
NAK 294	51	b	b	b	b	4
NAK 353	52	d	a	a	a	5
NAK 384	53	e	a	b	c	6
NAK 75	46	f	a	b	c	7
NAK 354	35	g	b	c	c	8
NAK 111	18	h	c	d	d	9

‡Underlined strains were chosen for 16S rRNA gene sequencing in subsequent step. Reference strains *R. leguminosarum* strain 8002 and *R. tropici* strain CIAT 899 are in bold

†RP01 subgroup from Figure 2.2

*Letters represent unique banding patterns on 1.5% agarose from digestion of 16S rRNA genes with the indicated endonucleases

#16S rRNA groupings numbered 1-9 represent the combination of restriction patterns obtained with the four endonucleases

nd- not determined

2.3.4 Phylogeny of strains based on 16S rRNA gene

A 1,292 bp alignment of sequenced 16S rRNA genes was analyzed and pairwise nucleotide identities within the 18 study strains ranged from 93.2% to 100%. Pairwise nucleotide identities of the sequences to those of their closest type strains of described species (all in the genus *Rhizobium*) ranged between 99.5% and 100% (Appendix 1).

A phylogeny based on the maximum likelihood algorithm placed the study strains into five clades (Figure 2.3). In clade A, five strains (NAK 210, NAK 266, NAK 312, NAK 334 and NAK 358) had 100% nucleotide identity with the type strains of *R. acidisoli*, *R. anhuiense*, *R. gallicum*, *R. laguerreae* and *R. sophorae*. NAK 382 was also part of group A, sharing 99.7 to 99.9% sequence identity with members of this cluster. Due to the high sequence identities observed between the study strains and multiple described species, study strains in this clade could not be assigned to any of the species by 16S rRNA alone.

Eight strains grouped into Clade B (NAK 287, NAK 239, NAK 103, NAK 245, NAK 157, NAK 458, NAK 35 and NAK 387), clustering with type strains of five rhizobial species (*R. phaseoli, R. fabae, R. pisi, R. etli* and *R. sophoriradicis*). The pairwise nucleotide identities within this group ranged between 99.5% and 100%. Minor sub-clades were observed, but these lineages were in most instances poorly supported by bootstrap values (<50%). Study strains falling into clade B could not be identified at the species level by 16S rRNA due to their high sequence identities to type strains of multiple species.

NAK 120 and NAK 349 shared a 100% nucleotide identity with the type strains of *R. jaguaris* and *R. paranaense* forming clade C. This group was closely related to clade D, in which NAK 294 had a nucleotide sequence 100% identical to that of *R. leucaenae* USDA 9039^T. The 16S rRNA gene of NAK 294 also contained the 72 bp insertion found in the 16S rRNA gene *R. leucaenae* USDA 9039^T (Ribeiro et al., 2012). These sequences were however not part of the alignment used for the phylogenetic reconstruction as their removal had been necessitated by the need to trim sequences of test and type strains to an equal length following alignment. NAK 384 was in clade E and was closely related to *R. pusense*. Study strains in clades C, D and E showed high sequence relatedness (>99.9%) to individual type strains of described rhizobial species, resulting in well

supported lineages (>70% bootstrap values). This suggested these strains could be identified at the species level by 16S rRNA alone.

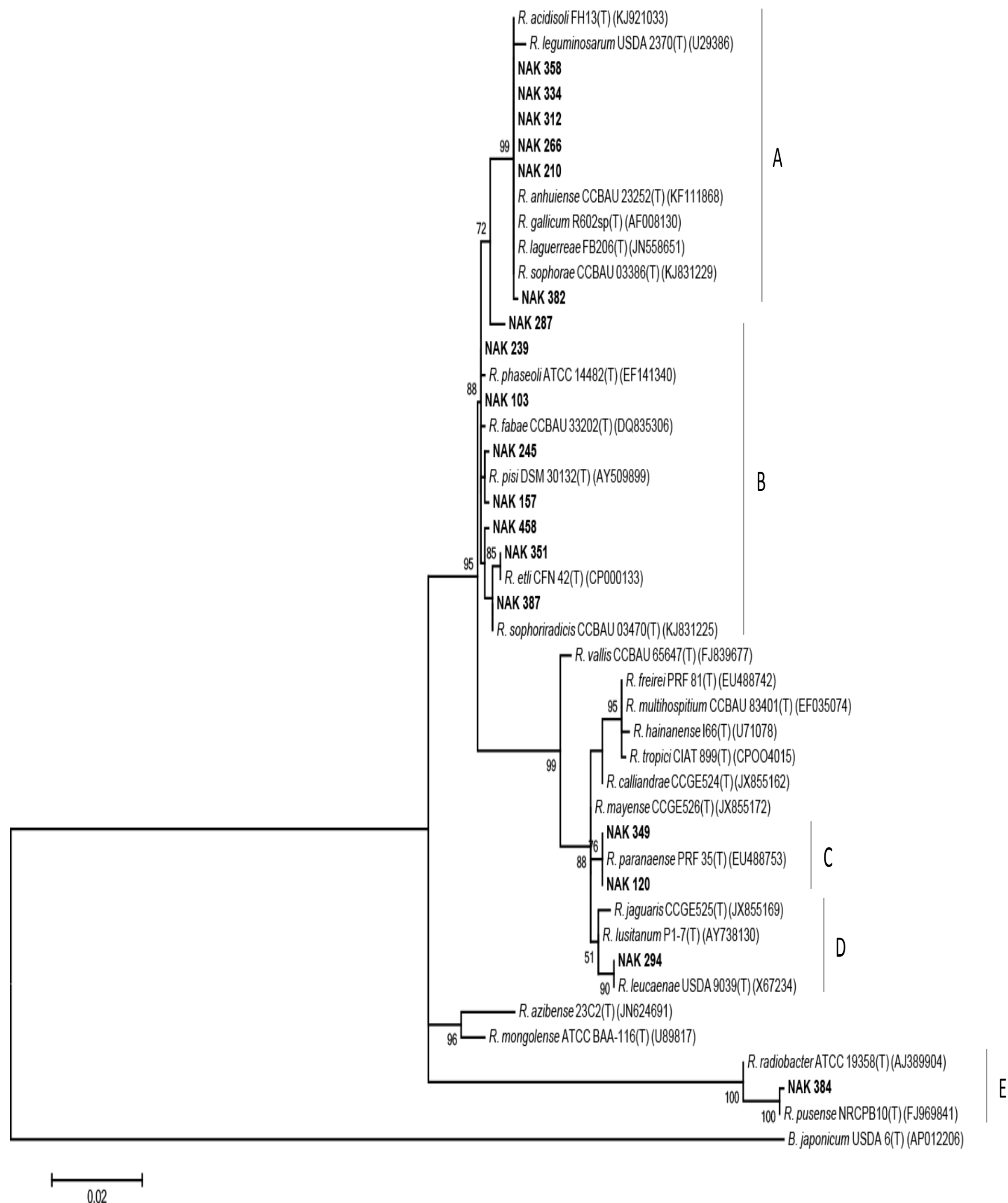


Figure 2.3: Maximum Likelihood phylogeny of the 16S rRNA gene showing the relationship between a subset of strains isolated in this study and their closest type strains. Accession numbers for type strain sequences are given after (T) denoting the type strains. There were a total of 1292 positions in the final dataset, and only node supports higher than 50% are labelled with a bootstrap value (500 replicates). The sequence of *Bradyrhizobium japonicum* USDA 6^T was included as an out group. Bar indicates 1 nucleotide substitutions per 100 nucleotides. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013).

2.3.5 Phylogeny of strains based on *recA* gene

To further explore the phylogeny of the 18 strains analyzed by 16S rRNA (Section 2.3.4), partial sequences of the *recA* gene (~500 bp) were obtained and aligned with *recA* sequences of the type strains of closely related rhizobial species following a BLAST search in GenBank. As the sequences of the type strains were considerably shorter than the sequences from this study, the final alignment was trimmed to 342 bp before analysis. The pairwise nucleotide identities within the study strains for this gene ranged from 82.4% to 100% while the pairwise nucleotide identities of the study strains to the sequences of their closest type strains of described species ranged between 93.6% and 100% (Appendix 2).

The phylogenetic analysis grouped the strains into seven clades (Figure 2.4). Clade A had a single strain, NAK 334, which had a 6.4% sequence divergence to the closest type strain and 5.4% divergence to the closest study strain. Five strains (NAK 210, NAK 266, NAK 312, NAK 358 and NAK 382) had nucleotide sequences that were 100% identical. These strains formed a well-supported monophyletic group (93% bootstrap value) that together with *R. acidisoli* FH13^{T}, which had a 96.6% identical nucleotide sequence, made up clade B.

Another five strains clustered with two type strains in two separate clades (C and D). The maximum sequence relatedness across members of the two clades was 95.6%. In clade C, the nucleotide sequence of NAK 387 matched that of *R. sophoriradicis* CCBAU 03470^{T} 100% while in clade D, the sequences of NAK 458, NAK 103 and NAK 239 were 100% identical to those of *R. phaseoli* ATCC 14482^{T}. Also in this clade was NAK 287, which had a 99.7% nucleotide sequence identity to the rest of the clade members. NAK 157 and NAK 245 had nucleotide sequence identities of 96.6% and 97.9% respectively to the type strain of *R. etli* in clade E.

NAK 294 grouped with *R. leucaenae* USDA 9039^{T} (97.9 % nucleotide identities) in one of the two subclades of clade F. The second sub-clade contained NAK 349 and NAK 120 that had 96.9% and 99.7% nucleotide identities to *R. paranaense* PRF 35, respectively. Members of the two sub-clades of F were separated from each other by a minimum of 4.1% sequence divergence. Clade G had NAK 384 grouping with *R. pusense* and *R. radiobacter* although sequences diverged by a minimum of 5.4%.

Overall, the *recA* sequence divergence of up to 17.6% among study strains and of up to 6.4% between study strains and described species was considerably greater than the maximum of 0.5% seen with 16S rRNA genes thereby allowing a deeper phylogenetic branching with a greater power to resolve genetic relationships.

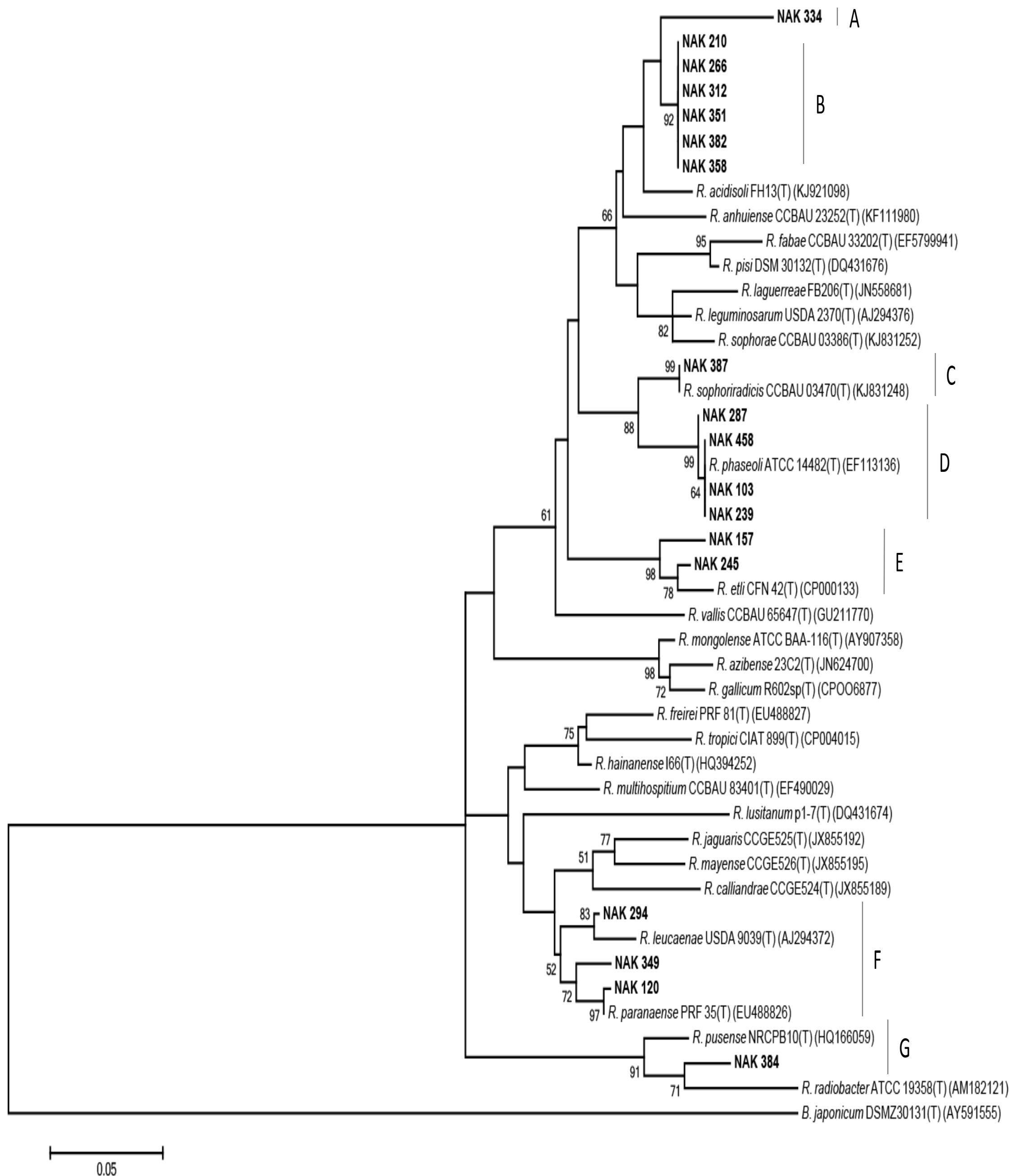


Figure 2.4: Maximum Likelihood phylogeny of the *recA* gene showing the relationship between a subset of strains isolated in this study and their closest type strains. There were a total of 342 positions in the final dataset, and only node supports higher than 50% are labelled with a bootstrap value (500 replicates). The sequence of *Bradyrhizobium japonicum* USDA 6^T was included as an out group. Bar indicates 5 nucleotide substitutions per 100 nucleotides. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013).

2.3.6 Phylogeny based on *nodC* genes

Partial *nodC* sequences were obtained from 15 of the 18 strains tested. No amplification was achieved with NAK 384, while poor quality sequence reads led to the omission of NAK 157 and NAK 351 sequences from the final analysis. The final alignment, following the addition of sequences from type strains and trimming to an equal length, was 489 bp. The pairwise nucleotide identities within the study strains for this gene ranged from 48.2% to 100%. Except NAK 120, the nucleotide sequences of each study strain matched the sequence of at least one type strain by 100%. NAK 120 had a sequence divergent from that of any type strain by a minimum of 3.7% (Appendix 3).

A phylogenetic analysis grouped the study strains into five clades (Figure 2.5). The corresponding *nodC* types were designated with Greek symbols (γ-a, γ-b, α, λ-a, λ-b) in line with previous reports (Aguilar et al., 2004; Rouhrazi et al., 2016). The strains NAK 312, NAK 334, NAK 266, NAK 245 and NAK 358 had nucleotide sequences that were 100% identical to the sequence of *R. vallis* CCBAU 65647^T and were in clade I (γ-a). Sequences of NAK 210, NAK 382 and NAK 387 were similarly 100% identical to those of *R. sophoriradicis* CCBAU 03470^T in clade II (γ-b). A 2.5% sequence divergence was seen between the two clades. However, the type strains of *R. sophorae* and *R. acidisoli* were positioned between the two clades and diverged from members of either clade by only 1.2% to 1.7%.

Four strains, NAK 103, NAK 239, NAK 287 and NAK 458 shared a 100% sequence identity with *R. etli* CFN 42^T and *R. phaseoli* ATCC 14482^T and formed clade III (α). Members of clade III were separated from those in clade II and clade I by a 3.5% and 3.9% sequence divergence, respectively.

NAK 120 was on an isolated, well supported (100% bootstrap) lineage in clade IV and the corresponding *nodC* type was designated λ-a. The sequence identity between NAK 120 and that of the closest type strains (of *R. tropici and R. leucaenae*) was 96.3% (3.7% divergence). Within clade V were NAK 294 and NAK 349. These two strains had nucleotide sequences that were 100% identical to those of *R. tropici* CIAT 899^{T}and *R. leucaenae* HBR12^{T}and 99.6% identical to *R. lusitanum* P1-7^T and their *nodC* was designated λ-b.

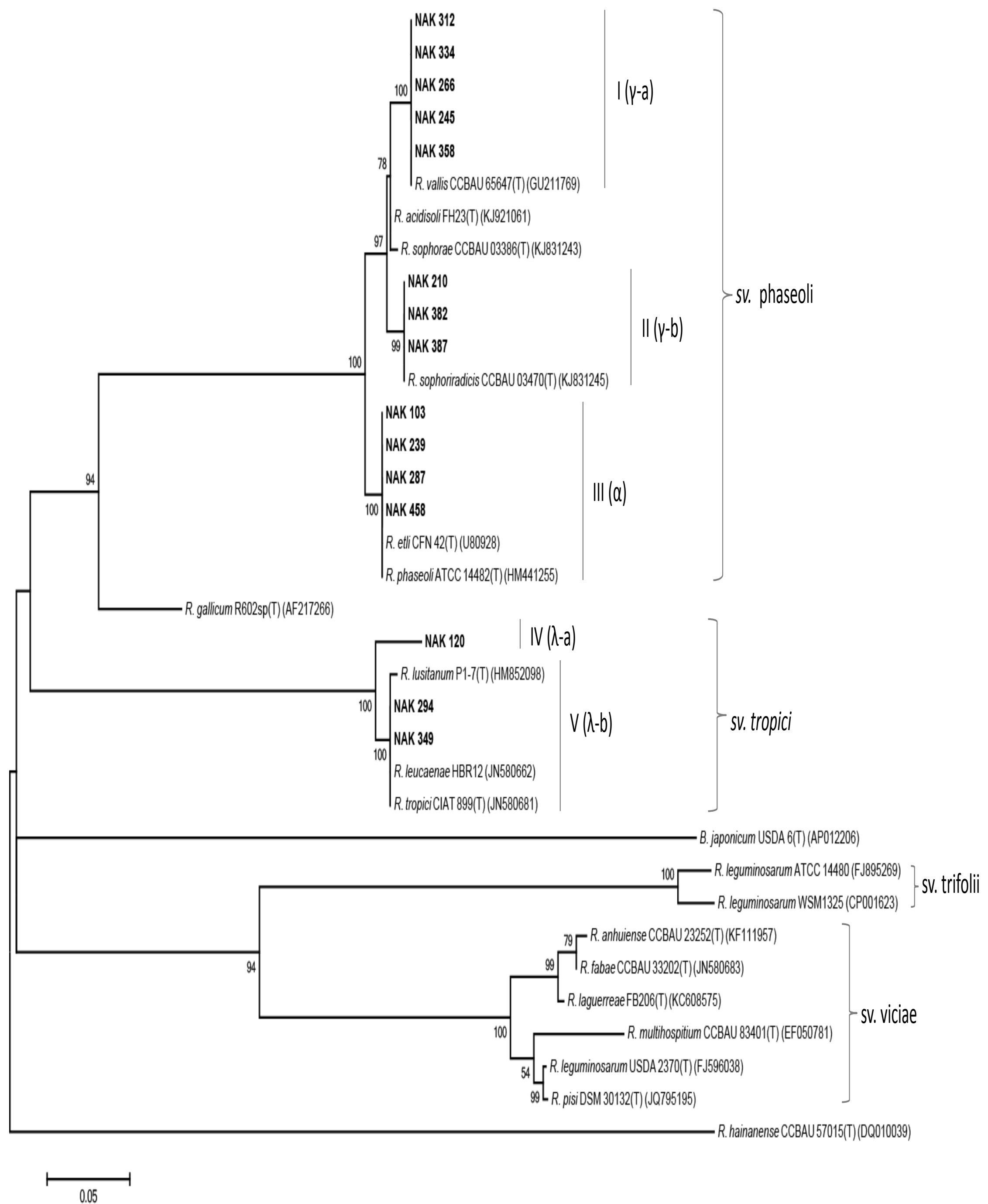


Figure 2.5: Maximum Likelihood phylogeny of the *nodC* gene showing the relationship between a subset of strains isolated in this study and their closest type strains. There were a total of 489 positions in the final dataset, and only node supports higher than 50% are labelled with a bootstrap value (500 replicates). The sequence of *Bradyrhizobium japonicum* USDA 6^T was included as an out group. Bar indicates 2 nucleotide substitutions per 100 nucleotides. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013). Four different *nodC* types are labelled I-IV.

2.4 Discussion

2.4.1 Diverse rhizobia nodulate *P. vulgaris* in Kenya

In a four-tiered reductionist approach, the genetic diversity of rhizobia collected was first assessed by DNA fingerprinting through RP01-PCR, revealing a large genetic diversity at the strain level. DNA fingerprinting by RP01-PCR has power for strain level discrimination (Collins et al., 2002; Denton et al., 2002; Thies et al., 2001) and strains showing greater than 80% similarity were regarded as clones for purposes of further analysis. An 80% cut-off was set conservatively below (a) the reported relatedness of banding patterns of clonal strains with common DNA fingerprinting techniques (Gevers et al., 2001) and (b) the 92% similarity of the banding profiles of duplicated NAK 73 RP01-PCR reactions in this study. Even with the use of the conservative 80% similarity cutoff to group strains, 88 RP01-PCR groups were obtained from 197 strains (Figure 2.2), indicating substantial genetic diversity among the strains.

The second major finding was that at least five current rhizobial species that can nodulate *P. vulgaris* occupy Kenyan soils. These are *R. sophoriradicis*, *R. phaseoli*, *R. leucaenae*, *R. paranaense* and *R. etli*. Strains were identified at the species level through the analyses of 16S rRNA and *recA* gene sequences following a preliminary estimation of the phylogeny of the rhizobial strains using PCR-RFLP analysis of 16S rRNA genes. An initial analysis of the relatedness of strains using RFLP of PCR-amplified 16 S rRNA genes has previously been used to analyze diversity in large rhizobia collections (Cardoso et al., 2012; Laguerre et al., 1996; Mnasri et al., 2007; Rouhrazi et al., 2016). The preliminary RFLP analysis in this study grouped 85 strains into nine groups (Table 2.4) and based on the representativeness of PCR-RFLP groups and sampling locations, 18 strains were selected for sequencing of the 16S rRNA and *recA* genes.

Phylogenetic reconstruction based on the 16S rRNA genes revealed that the majority of the sequenced study strains grouped with described species closely related to *R. leguminosarum* and *R. etli* (Clade A and B respectively in Figure 2.3), hereafter referred to as *R. leguminosarum* and *R. etli* lineages. The 16S rRNA genes of described species in the *R. leguminosarum* and *R. etli* lineages are highly conserved due to recent speciation (Lopez-Guerrero et al., 2012). For example, within the *R. leguminosarum* lineage, the type strains of *R. acidisoli*, *R. anhuiense*, *R. gallicum*, *R. laguerreae* and *R. sophorae* have 16S rRNA genes that are 100% identical (Aserse et al., 2012; Jiao et al., 2015; Ribeiro et al., 2015).

Consequently, the study strains clustering with members of either of the two lineages could not be identified based on the 16S rRNA genes. However, NAK 349, NAK 120 and NAK 294 (Clade C and D on Figure 2.3) were in the "*R. tropici* lineage" which displayed a well-supported phylogenetic branching. Consequently, NAK 349 and NAK 120 were identified as *R. paranaense* while NAK 294 was identified as belonging to *R. leucaenae*. NAK 384 (Clade E on Figure 2.3) was identified as belonging to *R. pusense*, a species originally described from the rhizosphere of chickpea (Panday et al., 2011). However, *nodC* was not successfully amplified from NAK 384 and further studies are needed to ascertain the strain's ability to nodulate *P. vulgaris*.

The phylogeny based on *recA* clarified the status of some of the study strains that could not be resolved by 16S rRNA genes. *recA* has previously, individually or in concatenation with other housekeeping genes, been used to produce robust phylogenetic trees that complement 16S rRNA phylogenies in rhizobia systematics (Aserse et al., 2012; Dall'Agnol et al., 2014; Gaunt et al., 2001; Martens et al., 2007). In the current study, 16S rRNA and *recA* sequences were found to be statistically incongruent ($P<0.05$). However, large sections of the resulting phylogenetic trees were found to have similar topologies (Figure 2.3 and Figure 2.4), indicating that only a few taxa were responsible for the incongruence. Greater sequence divergence of the *recA* genes allowed the identification of all strains falling within the *R. etli* lineage, except NAK 351. NAK 103, NAK 239 and NAK 287 belonged to *R. phaseoli*, NAK 387 to *R. sophoriradicis*, while NAK 157 and NAK 245 to *R. etli*. NAK 351 did not group consistently with the same strains on the two phylogenetic trees, and its taxonomic position was not apparent.

The analyses of the *recA* sequences of the strains falling within the *R. leguminosarum* lineage revealed that the *recA* sequences were highly divergent from those of any currently described rhizobial species. NAK 210, NAK 266, NAK 312, NAK 358 and NAK 382 had 100% identical *recA* nucleotide sequences amongst themselves but showed a maximum of 96.6% similarity in *recA* sequences to the type strain of any described species. The *recA* gene in NAK 334 had a 94.6% sequence similarity to the corresponding genes in the five strains above and a maximum of 93.6% similarity to the closest described species (*R. acidisoli*).

Given the divergent *recA* sequences of NAK 210, NAK 266, NAK 312, NAK 358 and NAK 382, it is likely that they belong to novel species. The minimum sequence

divergence between NAK 334 to that of a type strain of described rhizobial species was 6.4%, while the lowest divergence between NAK 210, NAK 266, NAK 312, NAK 358 and NAK 382 and a type strain was 3.4%. Of the currently described species in the genus *Rhizobium*, *recA* sequence divergence between species is as low as 2.7% (Appendix 2). Examples of *recA* sequence divergence between type strains of described species include: 2.7% between *R. fabae* and *R. pisi;* 2.8% between *R. sophorae* and *R. leguminosarum*; 2.8% between *R. gallicum* and *R. mongolense*; 3.4% between *R. gallicum* and *R. azibense*; 3.7% between *R. freirei* and *R. hainanense*; and 3.7% between *R. leguminosarum* and *R. laguerreae* (Appendix 2). Although the sequencing of more housekeeping genes is required to fully confirm the taxonomy of NAK 334, NAK 210, NAK 266, NAK 312, NAK 358 and NAK 382, their high *recA* sequence divergence from type strains of described species suggests they likely belong to two novel rhizobial species in the genus *Rhizobium*.

Profoundly, approximately 65% of the strains analyzed in this study might belong to putative novel taxa. All six strains likely to be of novel taxa belonged to PCR-RFLP group 1 as were 65% of the study strains (Table 2.4). Although the six strains were isolated from 6 separate sites in Nairobi, Kiambu and Meru (Table 2.3 and 2.4), the members of the wider PCR-RFLP group 1 were the predominant isolates from all sampling areas. By a similar analysis, 24% of the strains in this study belonged to species affiliated to the *R. etli* lineage, with *R. phaseoli*, *R. sophoriradicis*, and *R. etli* isolated. Sporadic *P. vulgaris* nodulators belonged to *R. paranaense* and *R. leucaenae*. An even greater diversity may exist in the rhizobial collection considering the conservative 80% cutoff used with RP01-PCR profiles and the insensitivity of the PCR-RFLP procedure. The two methods were utilized for the preliminary grouping of isolates.

The findings on the taxonomy of the rhizobia were largely in agreement with those of Anyango et al. (1995) who isolated strains resembling *R. etli*, *R. leguminosarum* and *R. tropici* from two locations in Kenya. Strains belonging to *R. etli* were isolated in this study. Although no *R. leguminosarum* or *R. tropici* were isolated in the current study, related species were recovered. For example, *R. paranaense* and *R. leucaenae,* closely related to *R. tropici,* were in this study found to nodulate *P. vulgaris* in Kenya. *R. leucaenae* and *R. paranaense* were until 2012 and 2014, respectively, classified as *R. tropici* (Dall'Agnol et al., 2014; Ribeiro et al., 2012). Additionally, a significant number of strains found to belong to putative new taxa were within the *R. leguminosarum* lineage

(Figure 2.3), and these may relate to the *R. leguminosarum* found nodulating *P. vulgaris* in Kenya by Anyango et al. (1995). Thus the current study updates the knowledge on rhizobia that nodulate *P. vulgaris* in Kenyan soils. It reports that strains potentially belonging to new taxa are major nodulators of *P. vulgaris* in Kenyan soils and that at least five species of *Rhizobium* also nodulate *P. vulgaris* in Kenya.

2.4.2 Rhizobia symbiovars nodulating *P. vulgaris* in Kenya

The analysis of the *nodC* genes revealed that the study strains were polymorphic for the gene, with four *nodC* types observed (Figure 2.5). The *nodC* gene is used to assign rhizobia to symbiovars (sv.) (Faghire et al., 2012; Rogel et al., 2011; Rouhrazi et al., 2016) and *nodC* types γ-a, γ-b and α placed strains in the sv. phaseoli while type λ-a and λ-b corresponded with sv. tropici (Figure 2.5). None of the study strains had *nodC* genes corresponding to sv. gallicum or sv. giardnii as has been reported elsewhere in strains from *P. vulgaris* in Ethiopia, France and Morocco (Armager et al., 1997; Aserse et al., 2012; Faghire et al., 2012).

The three alleles of the symbiovar phaseoli *nodC* seen in this study (γ-a, γ-b and α) have a wide global distribution among rhizobia nodulating *P. vulgaris* (Aguilar et al., 2004; Rouhrazi et al., 2016). For example, *R. vallis* CCBAU 65647^T isolated from China (Wang et al., 2011) has a *nodC* type γ-a, *R. sophoriradicis* from China (Jiao et al., 2015) has type γ-b, while type α sequences have been reported in strains from Mexico (Aguilar et al., 2004).

NAK 120 had *nodC* type λ-a, which was different (3.7% sequence divergence) (Figure 2.5; Appendix 3) from known strains in sv. tropici (Ormeño-Orrillo et al., 2012). NAK 294 and NAK 349 had *nodC* of type λ-b as did *R. tropici* CIAT 899^T isolated from Colombia (Martinez-Romero et al., 1991) and *R. leucaenae* HBR12 from Ethiopia (Aserse et al., 2012).

The *nodC* polymorphism among the strains has important implications on their host range. NodC proteins are determinants of chain lengths of lipo-chitoligosaccharides (LCOs) produced by rhizobia in response to flavonoids (Kamst et al., 1997). The length of the LCO significantly affects the structure of the LCO and hence its recognition by plant receptors and therefore host range (Oldroyd & Downie, 2008; Perret et al., 2000).

While it has been demonstrated that symbiovars are related to the host range (Rogel et al., 2011), differences in host range also exist within symbiovars. For example, *R. phaseoli* ATCC 14482^T nodulates *Trifolium repens* in addition to *P. vulgaris* (Ramirez-Bahena et al., 2008) while *R. vallis* CCBAU 65647^T and *R. sophoriradicis* CCBAU 03470^T nodulate *P. vulgaris* but not *Trifolium repens* (Jiao et al., 2015; Wang et al., 2011). The three strains are sv. phaseoli but show divergence in their *nodC* (α, γ-a, γ-b in corresponding order). In another report, *R. etli* CE3 nodulated *Lotus japonicus* while *R. etli* KIM5s did not (Pacios-Bras et al., 2002) in spite of both belonging to sv. phaseoli. A reasonable speculation is, therefore, that the various *nodC* types within sv. phaseoli seen in the current study are linked to the host range of the strains, allowing some of the strains to associate with alternative hosts in the absence of *P. vulgaris*.

Alternatively, the heterogeneity in *nodC* may relate to the *P. vulgaris* cultivars grown in Kenya. *P. vulgaris* of Mesoamerican origin has been reported to preferentially nodulate with strains carrying *nodC* type α in comparison to other *nodC* types (Aguilar et al., 2004). Kenyan farmers grow a mix of *P. vulgaris* from different centers of origin (Asfaw et al., 2009) and as no attempt was made during nodule sampling to identify cultivars, the effect of cultivars on the prevalence of *nodC* types will need to be explored in future studies. Members of the putative novel taxa nodulating *P. vulgaris* in Kenya had either *nodC* type γ-a or γ-b and these are likely the primary *nodC* types of rhizobia that nodulate *P. vulgaris* in the agro-ecological zones covered in the study. Strains belonging to species in the *R. etli* lineage had *nodC* type α apart from NAK 245 that had *nodC* type γ-a. Inter taxa similarity of *nodC* types is suggestive of lateral transfer of these genes between members of different taxa nodulating *P. vulgaris* in Kenya.

2.4.3 Concluding remarks

Kenyan soils surveyed harboured at least five current *Rhizobium* spp. that can nodulate *P. vulgaris*. In Central Kenya, the major group of strains nodulating *P. vulgaris* likely belong to putative novel species. Heterogeneity was seen in the *nodC* genes of the strains, and this might have important implications for the host range, survival and dominance of the strains. Additionally, this study demonstrated that rhizobia that associate with *P. vulgaris* in the centres of origin were present in Kenyan soils although only as minor nodulators. Approximately 65% of the bean symbionts recovered were genetically different by *recA* from any described species but, interestingly, were found to carry nodulation genes reported in rhizobia from other parts of the world. The results

presented here enhance existing knowledge on the distribution of *P. vulgaris* rhizobia and the prevalence of particular nodulation genes around the world.

The strains described in this chapter have not been vigorously tested for their capacity to fix N_2 and the next important step is to assess them for N_2 fixation. Some of the strains may form highly effective symbiosis with *P. vulgaris* and therefore be ideal inoculant strains for *P. vulgaris* in Kenya, due to a pre-adaption to Kenyan soil conditions.

CHAPTER 3
Effectiveness of strains on *P. vulgaris* under controlled conditions

3.1 Introduction

Rhizobial strains differ in their ability to fix N_2 (Section 1.2.3.1). Strains that fix large amounts of N can be applied as inoculants to increase the yield of host legumes, and, potentially, successive crops (Hungria et al., 2006; Peoples et al., 2009). However, inoculation of legumes does not guarantee yield gains and to increase benefits from inoculation, the inoculant strains need to be highly effective at N_2 fixation with the host, well-adapted to the edaphic conditions in the area of introduction and competitive for nodulation of legume against indigenous rhizobial strains (Gerding et al., 2014).

In Kenya, inoculation of *P. vulgaris* often does not result in any improvements in yields (Musandu & Ogendo, 2001; van der Bom, 2012). Many factors lead to the low success rate but a major one is the use of inoculant strains that do not meet the criteria of a good inoculant strain: effective, adapted and competitive. The main strain used for *P. vulgaris* inoculation in Kenya, *R. tropici* CIAT 899, is effective on Kenyan cultivars (Anyango et al., 1995; Kawaka et al., 2014) but is sometimes poorly adapted to the edaphic conditions in parts of the country. For example, CIAT 899 being an acid tolerant strain, has been shown to perform poorly in alkaline soils due to low adaptability (Anyango et al., 1998). In such soils, effective indigenous strains might be better suited as inoculants due to their greater adaptability to the soil conditions (Howieson & Ballard, 2004; Hungria et al., 2000). To identify potential inoculant strains for *P. vulgaris* from among strains indigenous to Kenyan soils that are pre-adapted to Kenyan soils, I assessed the effectiveness of strains from Chapter 2 on *P. vulgaris* in controlled glasshouse conditions. The effectiveness of strains was assessed through the comparison of shoot dry weights, N content and nodule scores of inoculated plants with those of the CIAT 899 treatment.

Leveraging on the effectiveness gradient present in the rhizobial collection, I also investigated the relationship between poly-β-hydroxybutyrate (PHB) accumulation in bacteroids, *nodC* alleles and strain effectiveness. In *P. vulgaris* symbiosis, the effect of poly-β-hydroxybutyrate (PHB) accumulation in bacteroids on N_2-fixation remains unclear. Previous studies using loss-of-function mutants unable to synthesize PHB have failed to elucidate the role of PHB in strain effectiveness (Cevallos et al., 1996; Lodwig et al., 2005) possibly due to the pleiotropic nature of PHB genes. PHB mutants are also defective in EPS production (Aneja et al., 2004; Trainer, 2009) and utilisation of various carbon sources (Cai et al., 2000; Cevallos et al., 1996; Lodwig et al., 2005),

characteristics that can alter symbiotic outcomes, independent of PHB accumulation. As an alternative to the use of mutants, I took advantage of the variability in N_2 fixation by strains (isolated in Chapter 2) to explore the association between strain effectiveness and PHB accumulation. To do this, PHB accumulation of variably effective strains was scored from electron micrographs and correlated with effectiveness.

Strains isolated from *P. vulgaris* in Kenya harboured at least five alleles of *nodC* that classified the strains into symbiovars phaseoli and tropici (Section 2.3.6). NodC is a determinant of host range in rhizobia (Perret et al., 2000) but it is not known if it is an indicator of strain effectiveness. To explore the relationship between *nodC* allele and effectiveness, I analysed the differences in effectiveness of strains with the various *nodC* alleles (Section 2.3.6).

Lastly, I investigated the effectiveness of the *P. vulgaris* symbiosis. Strains evaluated for N_2 fixation on legumes are identified as effective (or not) based on their performance against each other, a known reference strain, a N-free treatment, or against N-fed controls supplied with arbitrary amounts of N. As these approaches do not take into account the N requirements of the leguminous host, an "effective" strain may still not fix sufficient amounts of N to meet the N needs of the host. The ability of N_2 fixation, by rhizobia, to meet the N demands of *P. vulgaris* under controlled conditions is unclear. Consequently, I assessed the ability of an "effective" strain to meet the N requirements of *P. vulgaris*. The dry matter yield and N content of plants inoculated with the best strain from those tested was compared to that of plants receiving incremental amounts of mineral N to give an indication of the limits of N_2 fixation in *P. vulgaris*.

3.2 Materials and Methods

3.2.1 Bacterial strains and *P. vulgaris* cultivar

The studies in this chapter used a total of 54 rhizobial strains (Table 3.1). Fifty two of these represented the range of genetic diversity and geographical origin of the 197 strains isolated from Kenyan soils in Chapter 2 (Section 2.3.1), while two were reference strains. The reference strains were *R. tropici* CIAT 899, the leading commercial inoculant strain for *P. vulgaris* in Kenya, and *R. leguminosarum* 8002, a well-studied *P. vulgaris* strain. *P. vulgaris* cv. KK08, a determinately growing cultivar, was obtained from Fred Baijukya (N2Africa Project) and used as a host.

Table 3.1: A list of rhizobial strains used in this chapter

Kenyan strains*	Reference
NAK 69, NAK 75, NAK 91, NAK 103, NAK 104, NAK 105, NAK 111, NAK 120, NAK 157, NAK 210, NAK 214, NAK 220, NAK 223, NAK 227, NAK 231, NAK 239, NAK 242, NAK 245, NAK 254, NAK 266, NAK 270, NAK 284, NAK 287, NAK 288, NAK 294, NAK 295, NAK 299, NAK 303, NAK 312, NAK 315, NAK 321, NAK 327, NAK 332, NAK 334, NAK 343, NAK 349, NAK 351, NAK 353, NAK 354, NAK 358, NAK 363, NAK 367, NAK 368, NAK 378, NAK 382, NAK 387, NAK 403, NAK 405, NAK 407, NAK 440, NAK 441, NAK 458	This study
Reference strains	
R. tropici CIAT 899	Martinez-Romero et al. (1991)
R. leguminosarum 8002	Johnston et al. (1982)

*More details in Table 2.3

3.2.2 Assessment of N_2 fixation

Experiments to assess the effectiveness of strains were conducted in axenic vermiculite as described in Section 2.2.2, during the months of October to March in Perth, Western Australia. The study strains (Table 3.1) were assessed for N_2 fixation on *P. vulgaris* cv. KK08 in four batches. Each batch included a CIAT 899 and an un-inoculated, N-free treatment. The fourth batch re-evaluated 11 strains that represented the range of N_2 fixation observed in the initial three batches. Treatments in batches 1-3 were replicated in three pots, while those in batch 4 were replicated in five pots. Each pot contained three plants that were thinned to two plants 7 d after inoculation. Plants were harvested

42 d after inoculation, shoots excised and dried for 48 h at 60°C then weighed. At the time of shoot excision, roots were removed from pots and carefully washed free of adhering vermiculite and scored for nodulation using a nodulation score developed in preliminary experiments used to optimise glasshouse growth conditions for *P. vulgaris* (Appendix 4). For plants from the fourth batch, the percentage N in dry shoots was assessed as additional indicator of N_2 fixation. Dried shoots were pooled into two replicates, ground, and percentage N determined on a Leco F528 Nitrogen Analyzer at CSBP Soil and Plant Laboratory (Perth, Australia).

For data analysis, shoot dry weights (SDW) of inoculated plants were expressed as a percentage of the mean weight of the CIAT 899 treatment in their respective batch experiment and the resulting data compared across batches. Strains inducing weights ≥80%, 79%-50% 49-20% and <20% of CIAT 899 were as classified effective, partially effective, poorly effective and ineffective respectively. An analysis of variance (ANOVA) and Fisher's LSD (Section 3.2.5) was used to further compare strain effectiveness. Similarly, ANOVA and a post hoc test were used to compare percentage N in dry shoots from batch 4. Linear and nonlinear regression models were used to explore relationships between percentage N and SDW; total shoot N and SDW; and nodule scores and SDW.

3.2.3 Light and electron microscopy of nodules from variably effective strains

Light and electron microscopy was used to explore the anatomy of nodules from variably effective strains, and to quantify poly-β-hydroxybutyrate (PHB) accumulation in bacteroids of variably effective strains. To do this, representative nodules were selected from effective, partially effective, poorly effective and ineffective symbioses (Section 3.2.2) and light and electron microscope sections prepared as described by Spurr (1969) and Venable and Coggeshall (1965). Nodules were fixed overnight at 4°C in 3% (v/v) glutaraldehyde in 25 mM phosphate buffer (pH 7.0) before washing with a phosphate buffer and dehydration in a series of acetone solutions. Electron microscope material was post-fixed in 1% osmium tetroxide in 25 mM phosphate buffer (pH 7.0) for 2 h at room temperature before dehydration. Nodules were infiltrated with Spurr's epoxy resin and for light microscopy sectioned into 1-2 μM slices using a glass knife and sections stained with 1% (w/v) methylene blue and 1% (w/v) azurr II. Electron microscope sections were cut at approximately 90 nm, mounted on copper grids before double

staining in aqueous uranyl acetate and lead acetate. Light microscope examination was under an Olympus BX51 photomicroscope and photographs taken with an Olympus DP-70 camera while electron microscopy was on a Phillips CM100 Bio Twin Transmission Electron Microscope.

Five bacteroids were chosen from representative electron micrographs of each strain and used to estimate the amount of PHB accumulated in each bacteroid by the 'counting squares method'. To do this, a square grid was created over an electron micrograph and the proportion (of total) of squares covering sections with PHB in each bacteroid determined. For squares partially overlaying areas with PHB, squares with PHB greater than half of the square area were counted as whole while those with less than half were ignored. The resulting data were then analysed by ANOVA and Tukey HSD (Section 3.2.5), and PHB accumulation compared with effectiveness.

3.2.4 Comparing the growth of inoculated *P. vulgaris* cv. KK08 to that supplied with mineral N

The growth response of *P. vulgaris* cv. KK08 was assessed using the vermiculite system (Section 2.2.2) across a range of N-feed rates given as 0, 16.8, 33.6, 67.2, 100.8, 134.5, 168.1, 201.7 and 235.3 mg of N as NH_4NO_3 pot^{-1} week^{-1}. In addition to the nine N treatments, one treatment was inoculated with CIAT 899 at sowing. All treatments received 100 mL of 10 mM NH_4NO_3 as starter N 3 d after planting. Mineral N treatments were commenced 10 d after planting. A 1 M solution of NH_4NO_3 was prepared and aliquots applied twice a week together with other nutrients to achieve the desired N application rate. All treatments were replicated four times in pots containing three plants, later thinned to two. Plants were harvested 35 d post-planting and shoots were dried and weighed before determination of the N content. All other experimental procedures were as described in Section 2.2.2 and 3.2.2.

3.2.5 Data analysis

Where applicable, data were subjected to an analysis of variance (ANOVA) using SPSS version 22 (IBM Corp, released 2013). ANOVA was preceded by a test for normality and equal variances (Levene's test). Fisher's LSD was then used when ANOVA was found to be significant.

3.3 Results

3.3.1 Biomass and N accumulation in shoots of inoculated plants

The N_2 fixation effectiveness of 54 strains (including CIAT 899 and 8002) was assessed on *P. vulgaris* cv. KK08 by a combination of visual observation of plant vigour and leaf colour, assessment of nodulation, and an analysis of shoot dry weights. Un-inoculated plants were visibly chlorotic and lacked nodules, while inoculated plants were of variable leaf colour and vigour, and were nodulated (Figure 3.1).

Figure 3.1: A, *P. vulgaris* cv. KK08 plants in the glasshouse showing differences in growth vigour and leaf colour, following inoculation with different rhizobial strains. **B,** a nodulated *P. vulgaris* root system.

Shoot weights were expressed as a percentage of the weight of the CIAT 899 treatment in their respective batch experiment (Figure 3.2) to allow comparisons. CIAT 899 treatments had the highest mean shoot biomass accumulation, but post hoc comparisons using Fisher's LSD indicated 11 strains isolated from Kenya (NAK 407, NAK 458, NAK 354, NAK 327, NAK 227, NAK 214, NAK 104, NAK 288, NAK 239, NAK 157 and NAK 299) induced biomass comparable to CIAT 899 ($P>0.05$). Five strains (NAK 210, NAK 75, NAK 351, NAK 223 and NAK 69) did not result in dry matter yield greater than that of the un-inoculated control (LSD, $P>0.05$), although they induced nodules. There was a 12.5-fold difference in shoot dry matter between the

most effective symbiosis (CIAT 899) and the least (NAK 210), indicating a considerable variability in the effectiveness of the strains.

Strains were grouped into three effectiveness levels based on plant dry weights. Strains inducing weights ≥80%, 79%-50% 49-20% and <20% of CIAT 899 were classified effective, partially effective, poorly effective and ineffective respectively (Figure 3.2).

Twelve strains, including CIAT 899, were selected representing the range of effectiveness observed and in a subsequent experiment, tested for N accumulation in the shoots. Effective strains in this subset were CIAT 899, NAK 458, NAK 239, NAK 103 and NAK 287. Strains NAK 294, 8002, NAK 245, and NAK 334 were partially effective, NAK 312 and NAK 120 were poorly effective, while NAK 210 was ineffective. The lowest N concentration among the inoculated, of 3.17%, was found in the shoots of plants inoculated with NAK 210 while the highest N concentration, 4.5%, was from inoculation with NAK 120 (Figure 3.3).

Plotting shoot N concentration against shoot dry weight of treatments revealed that shoot dry weight increased with increase in shoot N concentration until an apparent optimum of between 3.7%-4.1% nitrogen was reached (Figure 3.3). Two treatments, NAK 334 and NAK 120, displayed N concentrations higher than this optimum coupled with substantially reduced dry matter. A non-linear regression function provided the best fit for the dataset (R^2=0.743) (Figure 3.3). As an example to illustrate the non-linearity of SDW vs. %N, CIAT 899 induced the greatest biomass accumulation but shoots contained only 3.73% nitrogen. On the other hand, plants inoculated with NAK 120 accumulated 42% of shoot matter accumulated by CIAT 899 plants but their N concentration was significantly higher at 4.5% (LSD, $P<0.05$).

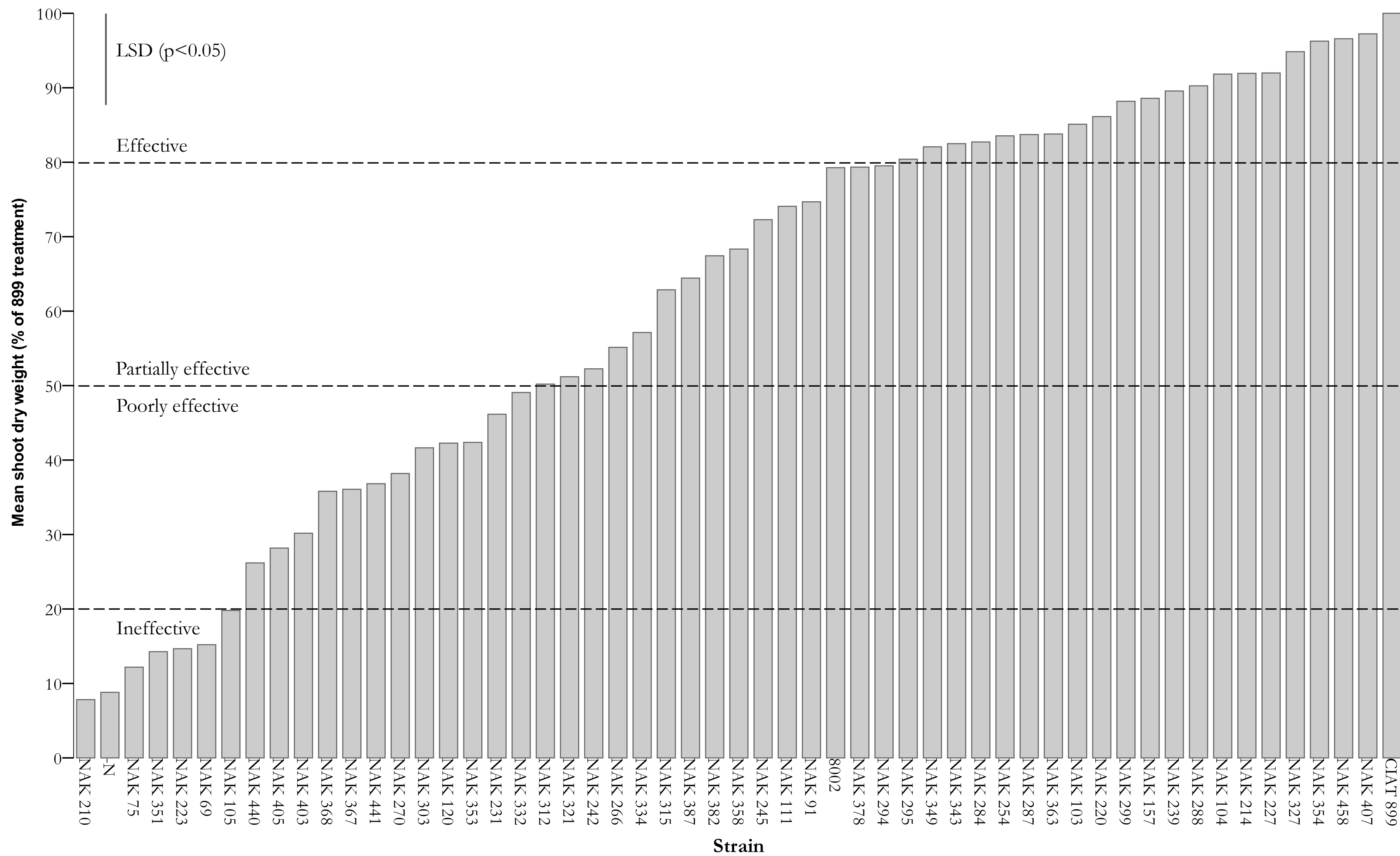


Figure 3.2: Mean shoot dry weights of *P. vulgaris* cv. KK08 inoculated with 54 different rhizobia strains expressed as a percentage of CIAT 899 treatment.

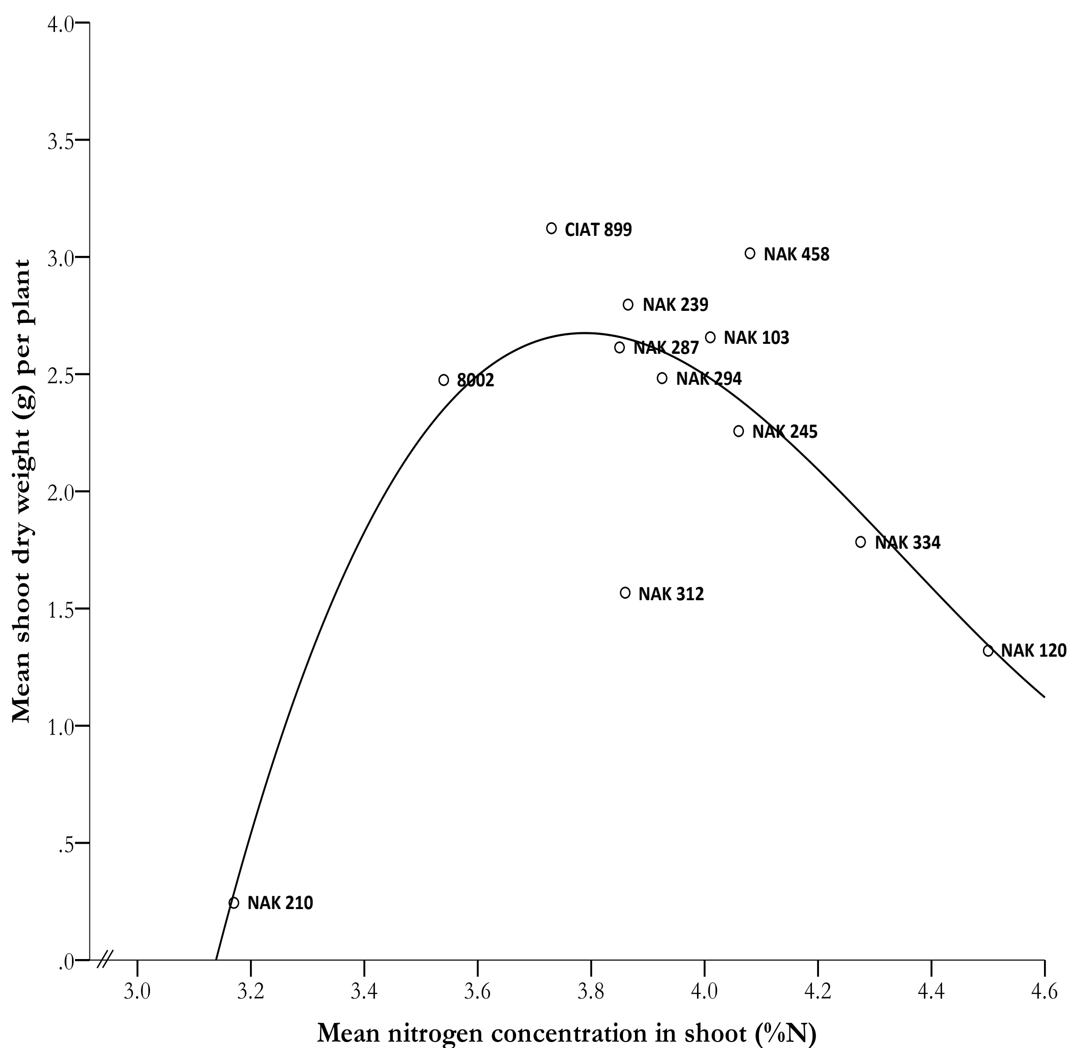


Figure 3.3: Plot of nitrogen concentration (%N) and shoot dry weight (SDW) of inoculated *P. vulgaris* cv. KK08 plants fitted using a nonlinear regression model (solid line). R^2=0.743. Data points are means of six SDW and two % N measurements.

Total N in shoots was greatest in plants inoculated with NAK 458 (123.22 mg) although the amount was not statistically different (P>0.05) from that in CIAT 899-inoculated plants. Shoots of plants inoculated with NAK 210 plants accumulated the least amount of N (7.61 mg). Unlike N concentration, total shoot N was positively and linearly correlated with shoot dry weight (R^2=0.976) (Figure 3.4).

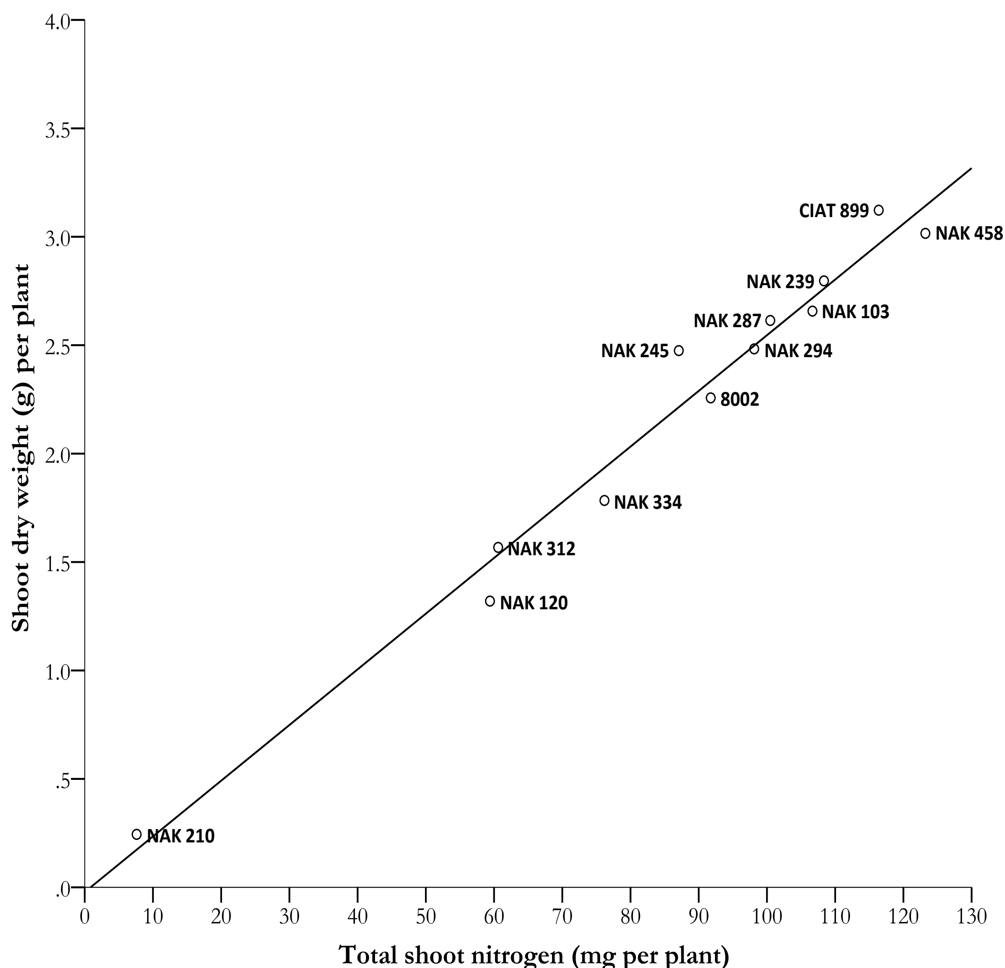


Figure 3.4: Plot of total shoot nitrogen (%N × SDW) and shoot dry weight (SDW) of inoculated *P. vulgaris* cv. KK08 plants fitted using a linear regression model (solid line). R^2=0.976. Data points are means.

3.3.2 Nodulation in *P. vulgaris*

Four main types of nodules were observed following the inoculation of plants with the different strains (Figure 3.5): pink nodules that varied in size from greater than 4 mm to less than 2 mm; white nodules, typically less than 2 mm in size; green nodules; and atypical nodules, such as those with extensive lenticels or tapered in shape.

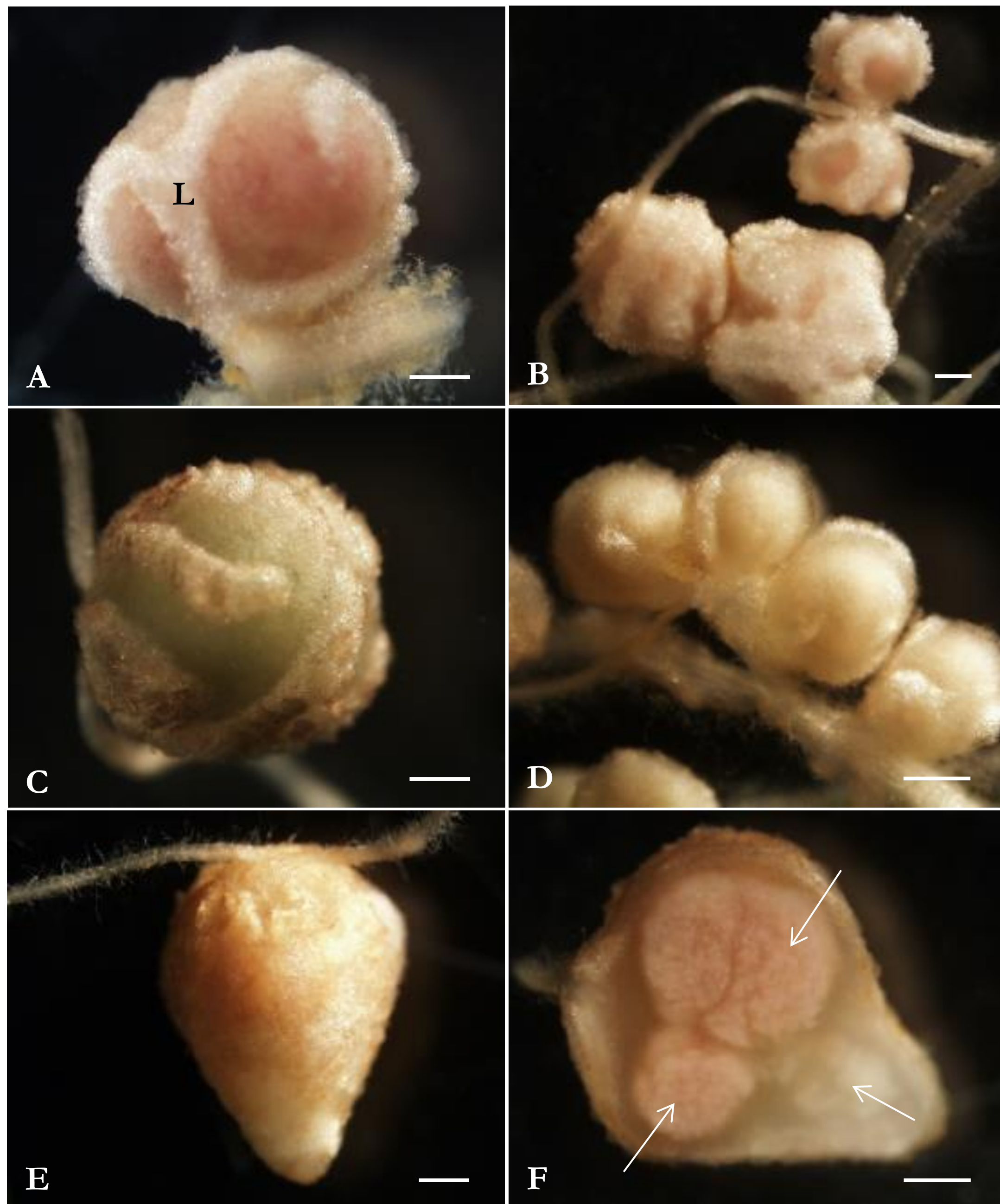


Figure 3.5: *P. vulgaris* root nodules. **A**: typical fixing desmodioid *P. vulgaris* nodule (from inoculation with CIAT 899). This was the most commonly observed type, and nodules exhibited the characteristic pink hue of fixing nodules and varied in number and size (range given in Appendix 4). Lenticels (marked L) are visible on the outer surface. **B**: nodules formed by NAK 287 showing a more extensive lenticel development. **C**: senescing nodule following inoculation with CIAT 899. The pink hue is replaced by a bright green pigment. **D**: partially effective nodules formed by NAK 210. Nodules are small, approximately 1 mm in diameter and the pink hue is not apparent. **E**: nodule with a distinctively smooth surface and lacking lenticels resulting from inoculation with NAK 120. The nodule is tapered with a growing tip. **F**: hand section of the nodule in E. The interior contains a tri-lobed infected zone (arrows). The lobe towards the tapered end is less developed. Bar is 500 µm.

The nodulation patterns, in general, differed in four ways: in nodule colour, nodule size, nodule number, and nodule distribution on the root system. These factors had significant interactions. For example, large (>4mm) pink nodules tended to be fewer (<100), and in the case of CIAT 899, were mostly in the crown of the root. On the other hand, NAK 210 formed numerous (>200) tiny (<1mm) white nodules on the laterals.

Nodulation was assessed using the four the 4-point scale developed in a preliminary experiment (Appendix 4) and nodulation was found to be positively correlated with plant shoot dry weight (R^2 of 0.87).

3.3.3 Anatomy of nodules formed by variably effective strains

Nodules from *P. vulgaris* inoculated separately with seven effective strains (CIAT 899, NAK 104, NAK 288, NAK 239, NAK 157, NAK 220, NAK 287), five partially effective strains (NAK 294, 8002, NAK 245, NAK 387, NAK 315, NAK 334), one poorly effective strain (NAK 312) and one ineffective strain (NAK 210) were sectioned and examined under a light microscope (Figure 3.6).

Nodules examined had similar internal organization consisting of a kidney-shaped central tissue surrounded by peripheral tissues. Peripheral tissues, from the outside to the inside included lenticels, outer cortex, endodermis and inner cortex (parenchyma) containing vascular bundles. The central region in all nodules contained a mixture of infected and uninfected host cells with variations in the relative abundance of each type. Nodules formed by effective strains had a higher proportion of cells in the central zone infected when compared to uninfected (Figure 3.6). On the other hand, nodules formed by ineffective strains had infected cells interspaced with a high proportion of uninfected cells (Figure 3.6). Additionally, nodules formed by ineffective strains sometimes had small central infection zones when considered as fractions of the entire nodule (Figure 3.6). Nodules formed by partially effective strains were not always readily distinguishable from those by effective strains based on the proportion of infected to uninfected cells (Figure 3.6B vs 3.6C).

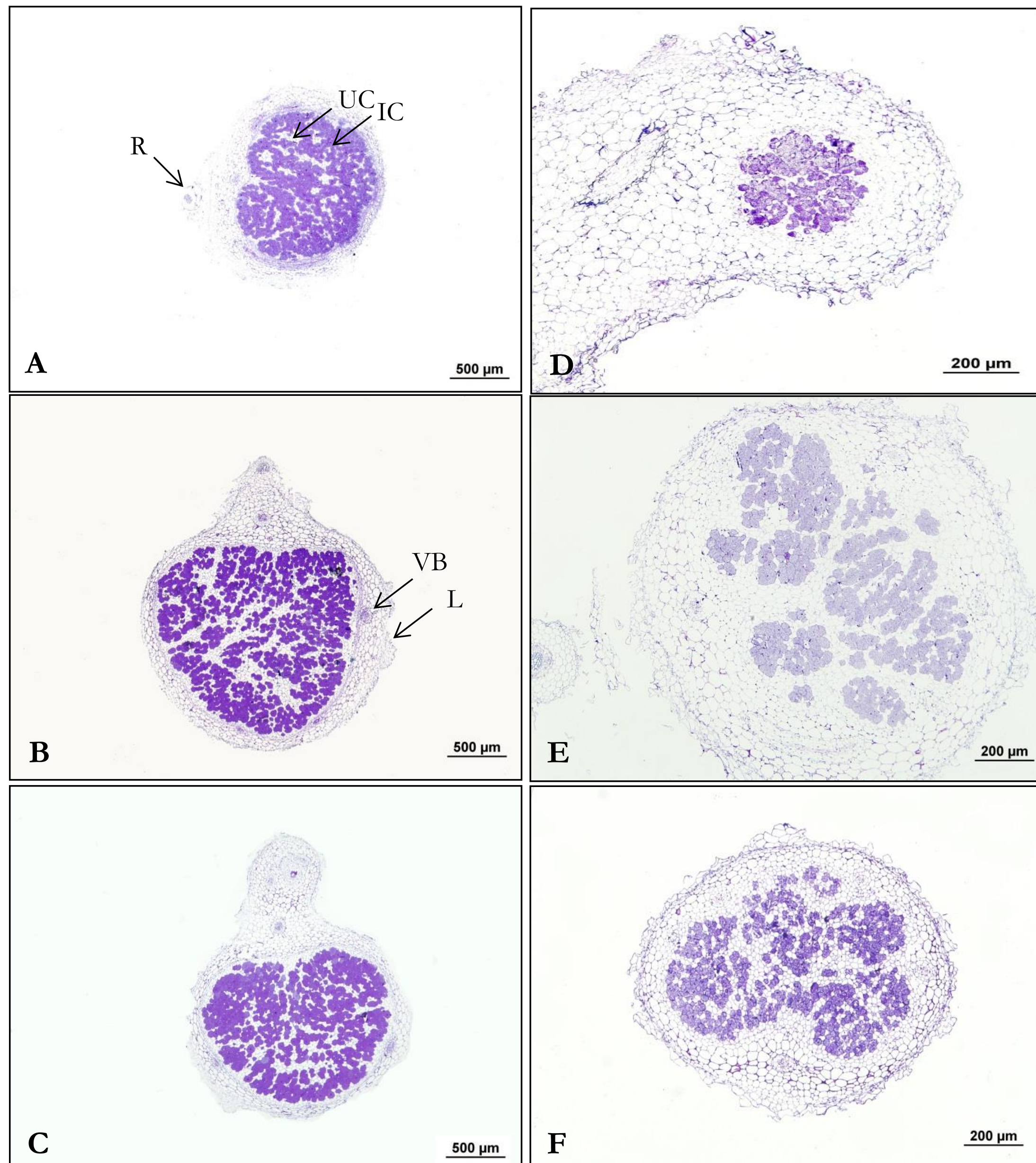


Figure 3.6: Light micrographs of root nodules (traverse to the primary root axis) from inoculation of *P. vulgaris* cv. KK08 with 6 different strains. **A**: NAK 239 (effective). **B**: NAK 157(effective). **C**: NAK 245 (partially effective). **D**: NAK 334 (partially effective). **E**: NAK 312 (poorly effective). **F**: NAK 210 (ineffective). Nodule sections A, B and C contain a densely infected central zone, while D has a high percentage of infected cells in a relatively small infection area. E and F contain a central zone with a markedly large proportion of uninfected cells. IC=infected central zone, UC=uninfected cells, R=root, VB=vascular bundle, L=lenticel.

3.3.4 Poly-hydroxybutyrate (PHB) accumulation in bacteroids

Nodules from *P. vulgaris* inoculated with nine variably effective strains were examined under an electron microscope (Figure 3.7). Effective strains were CIAT 899, NAK 354 and NAK 287. Partially effective were NAK 294, 8002 and NAK 245 while NAK 312,

NAK 353 and NAK 120 represented poorly effective strains. Electron micrographs revealed two or more rod-shaped bacteroids were often contained in each symbiosome. In a few instances, pleomorphic bacteroids were observed (3.8C). All nodules contained bacteroids with white/clear zones indicative of accumulation of the carbon storage compound, PHB. CIAT 899 bacteroids accumulated the least amount of PHB (21% of total cell area). All other eight strains resulted in bacteroids with PHB as a high proportion of cellular area (71%-83%) and differences were not significant (HSD, $P>0.05$).

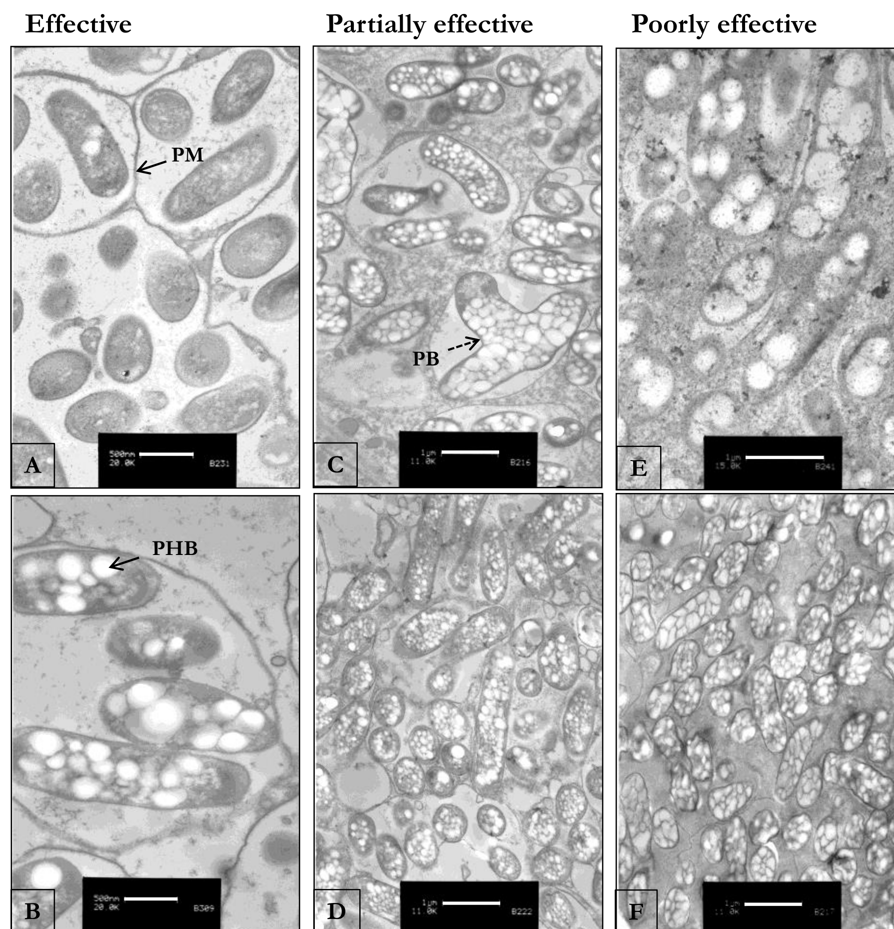


Figure 3.7: Sample transmission electron micrographs of nodules of *P. vulgaris* inoculated separately with six variably-effective strains of rhizobia, showing accumulation of poly-β-hydroxybutyrate (PHB) in bacteroids. Inoculation was done with **A**, *R. tropici* CIAT 899. The arrow indicates a peribacteroid membrane (PM) encapsulating several bacteroids that are primarily devoid of PHB granules. **B**, NAK 354 (PHB labelled shown with an arrow). **C**, NAK 245. Pleomorphic bacteroid is shown (arrow labelled PB). **D**, 8002. **E**, NAK 120 **F**, NAK 353. Bacteroids were 1-3 μm in length. Bar in **A** and **B** is 500 nm, and 1 μm for all other. Nodules had been harvested 42 d after inoculation.

3.3.5 Link between *nodC* allele and effectiveness

Fourteen of the 15 strains assessed for the *nodC* type (Section 2.3.6) were included in an analysis to investigate possible associations between *nodC* type and strain effectiveness. NAK 120 did not share a *nodC* type (λ-a) with any other strain and was, therefore, for statistical reasons excluded from the analysis. Each of the four remaining *nodC* types (γ-a, γ-b, α, and λ-b) was represented by data from 2-5 strains each, inoculated on six plants. The mean shoot dry matter accumulated by the strains harbouring the four *nodC* types, was expressed as a percentage of that of CIAT 899 (Figure 3.8). The mean SDW of strains with *nodC* type γ-a was 46.6%, type γ-b was 62.1%, type λ-b was 80.6%, and type α was 88.2%. An analysis of variance (ANOVA) revealed significant variations in the effectiveness of strains based on *nodC* alleles, $F (4, 11) = 15.276$, $p=0.000$. A post hoc Tukey test revealed strains with *nodC* type γ-a and γ-b; γ-b and λ-a; α and λ-b did not differ significantly ($p>0.05$) in their effectiveness. Differences between others were significant ($p<0.05$).

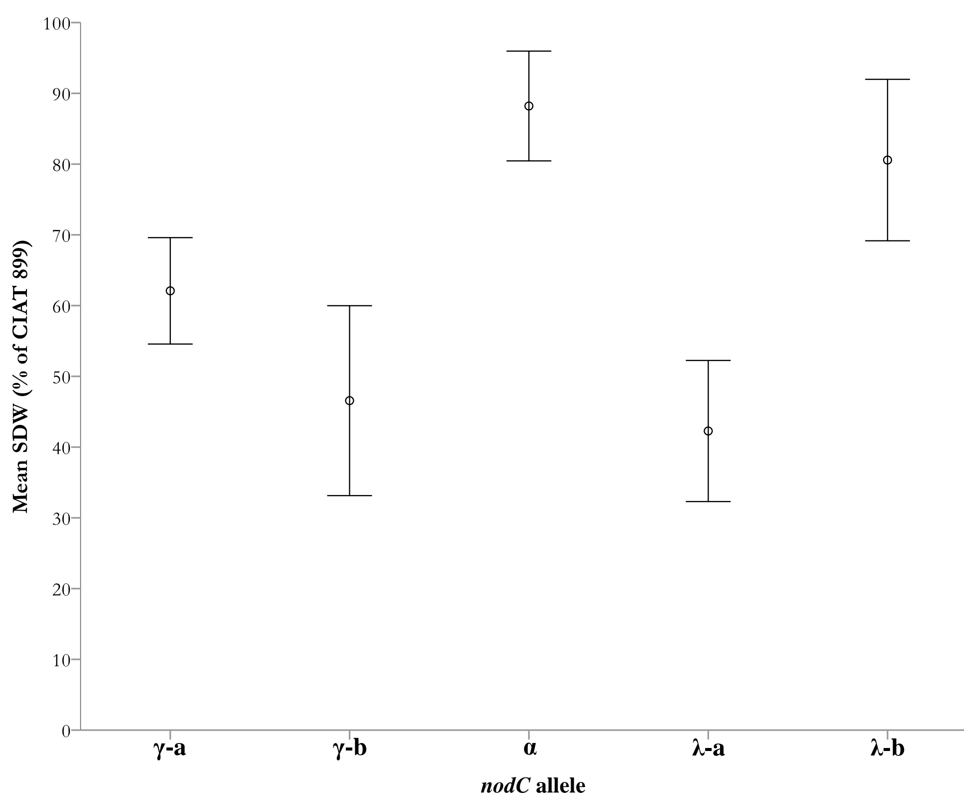


Figure 3.8: SDW of strains (as a percentage of CIAT 899) plotted against their *nodC* type. Error bars: 95% CI.

3.3.6 Growth of *P. vulgaris* under mineral N in comparison to BNF

An increase in the amount of N applied as NH_4NO_3 was strongly correlated with plant shoot mass (Pearson's r=0.9) (Figure 3.9, 3.10). The relationship was curvilinear with a steeper sigmoidal slope at lower N rates indicating greater growth response to N, followed by decreased response to N application at higher rates. The highest mean shoot mass was achieved at 201 mg N pot^{-1} $week^{-1}$. However, Fisher's LSD indicated no significant differences ($p>0.05$) between treatments receiving 134-235 mg N pot^{-1} $week^{-1}$. The highest shoot dry matter achieved per plant under N fertilization (3.04 g) was 2.5 times that of the CIAT 899 treatment (1.21 g).

Figure 3.9: Differences in the vigour of *P. vulgaris* cv. KK08 plants supplied with increasing amounts of N in the form of NH_4NO_3. The green arrow shows the direction of increasing N rate. CIAT 899 treatment is indicated on the right.

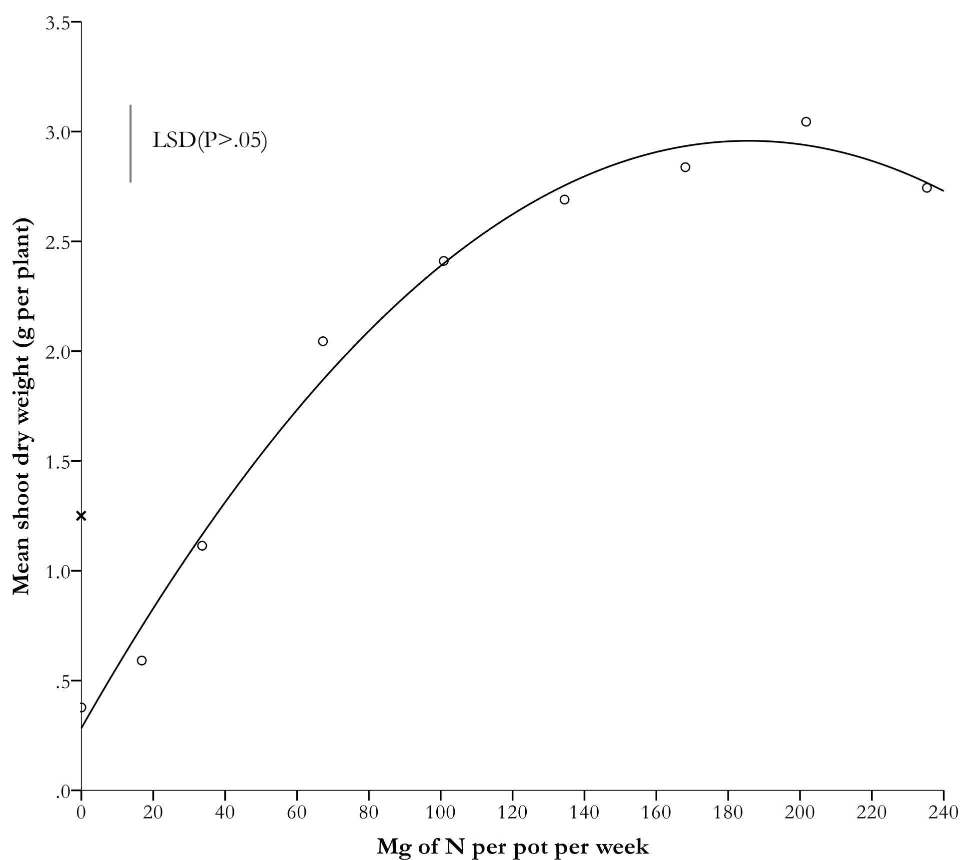


Figure 3.10: Mean shoot dry weights (SDW) of *P. vulgaris* cv. KK08 plants receiving either N as NH_4NO_3 (o) or inoculated with CIAT 899 (×) taken after 35 d of growth. Data is the mean of six plants. Data fitted using a nonlinear regression model. R^2 =0.991.

The shoot N concentration (%N) of plants receiving mineral N showed a linear relationship with their dry matter content (R^2=0.979). However, the SDW and % N measurements of CIAT 899 did not fit into this regression line. For the same concentration of N, CIAT 899 had significantly less biomass. The mean SDW in inoculated plants containing 4% N was 1.21g, while at a similar N concentration (4%), plants supplied with mineral N had a mean SDW of 2.05 g (Figure 3.11).

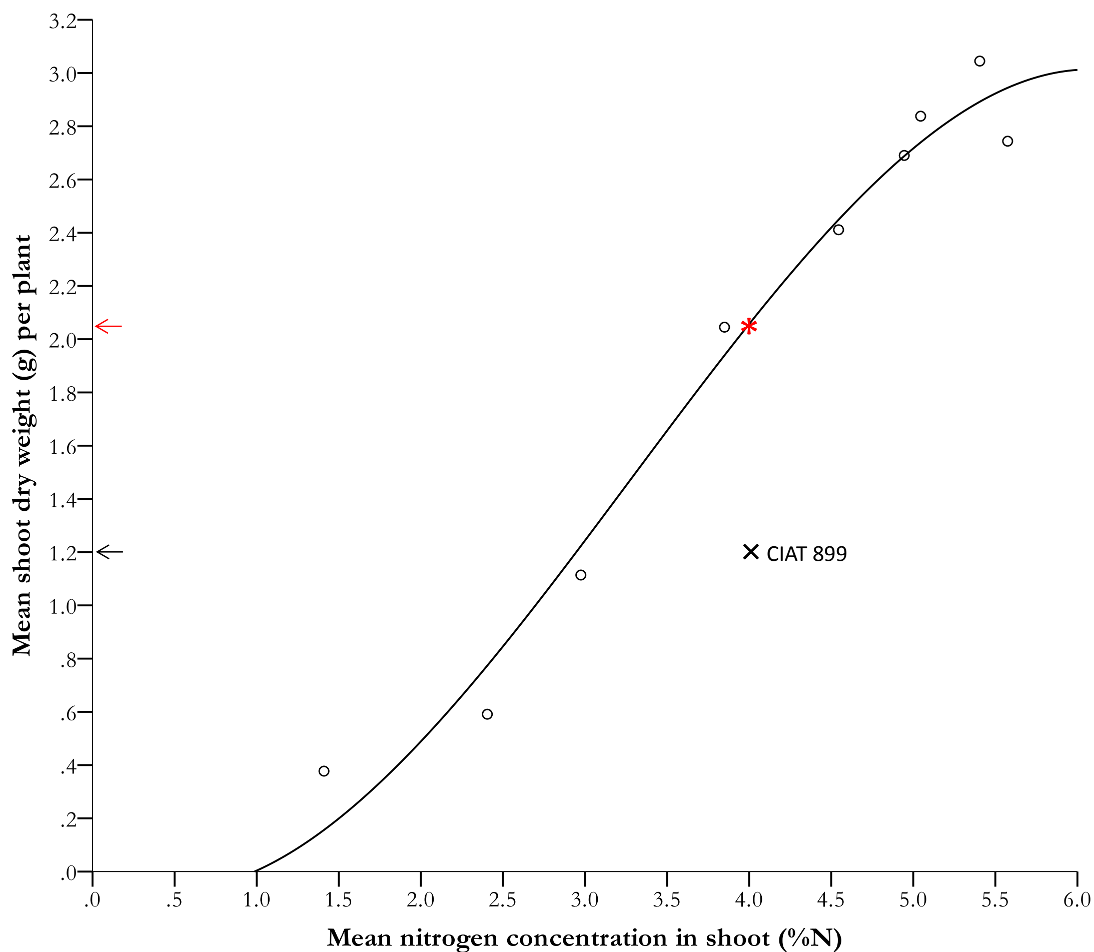


Figure 3.11: Percentage shoot N concentration graphed against SDW in *P. vulgaris* cv. KK08 plants receiving increasing rates of N as NH_4NO_3. Experimental SDW of CIAT 899 treatment (×) and expected based on N concentration (*) are shown. Plants were harvested 35 d after planting. Data fitted using a nonlinear regression model. R^2 =0.979.

3.4 Discussion

Studies in this chapter investigated the effectiveness of strains isolated from Kenyan soils on *P. vulgaris* and explored relationships between strain effectiveness and (a) anatomical nodule features, (b) PHB accumulation in bacteroids, and (c) *nodC* alleles harboured. Further, the studies evaluated the ability of the most effective strain to meet the N requirements of the host.

3.4.1 The effectiveness of strains

The 52 strains were found to be variably effective on *P. vulgaris* cv. KK08 (Figure 3.2). Importantly, the experiments uncovered highly effective strains which were as effective at N_2 fixation as CIAT 899, a leading inoculant strain for *P. vulgaris*.

To assess N_2 fixation by each of the strains, the shoot dry weights of inoculated plants were compared to the shoot dry weight of plants inoculated with CIAT 899. The dry weights of shoots are a common proxy for N_2 fixation in systems deficient in N, such as the vermiculite system used in this study (Howieson et al., 2005). When N is the only growth limiting factor, the growth of plants is positively correlated with the availability of N (Andrews et al., 1999; Field & Mooney, 1986). Of the 52 strains analysed for N_2 fixation, 20 were characterized as effective, 13 as partially effective, 13 as poorly effective and seven as ineffective. The delineations were arbitrary but chosen to give reasonable categories of strain effectiveness in relation to CIAT 899, in line with previous categorisations of effectiveness (Howieson et al., 2005). Eleven of the effective strains were statistically equivalent to CIAT 899 at N_2 fixation (P=0.05).

A very high prevalence of effectiveness was observed among the strains collected for this study. Although Kawaka et al. (2014) and Anyango et al. (1995) reported a high proportion of effective rhizobial strains from among isolates from Kenya, the high percentage of effective strains seen in the current study may have resulted from the nodule sampling strategy used. The sampling targeted healthy bean plants and may therefore have skewed the effectiveness spectrum of the isolated strains towards the effective end. Consequently, the proportions of effectiveness observed in this study may not be reflective of the proportions of effective of strains in the soils in areas of nodule sampling.

The 11 strains identified as being equal to CIAT 899 at N_2 fixation are candidate strains for field evaluation and possibly represent adapted and highly effective inoculant strains

for *P. vulgaris* in Kenya. The candidate strains were isolated from diverse agro-ecological zones differing in soil types, rainfall, and temperatures (Chapter 2, Table 2.2) and are consequently expected to display adaptability to specific environmental stresses present in those areas. CIAT 899 is the strain of choice for inoculation of *P. vulgaris* in Kenya (Bala et al., 2011). The strain possesses some excellent characteristics such as high rates of N_2 fixation, tolerance to environmental stresses such as soil acidity (Graham et al., 1994) and high temperatures (Martinez-Romero et al., 1991), and is genetically stable under environmental stresses (Hungria & Vargas, 2000). However, with recent reports of its failure to improve *P. vulgaris* yields in Kenyan fields (Gicharu et al., 2013; van der Bom, 2012), a new set of strains like the ones isolated in this study will need to be tested for suitability as inoculants.

3.4.2 Anatomical features of nodules

This study characterized a large number of strains for effectiveness at fixing N_2 with *P. vulgaris* resulting in a N_2 fixation gradient (Figure 3.2) spanning the entire range expected (Howieson et al., 1995). Strains at different points in the N_2 fixation gradient then provided an excellent opportunity to investigate key anatomical features of nodules induced by variably effective strains

Effective strains formed nodules with a greater proportion of cells in the central zone infected with bacteroids in comparison to nodules formed by ineffective strains (Figure 3.6). A mix of infected and uninfected cells in the central tissue of nodules is commonly found in phaseolid legumes (Sprent et al., 2013; Tate et al., 1994) in contrast to legumes such as *Listia* and *Lotononis* that contain a uniformly infected central tissue (Ardley et al., 2013). The specific events leading to differences in proportions of infected and uninfected cells remain unclear but would be related to differences in the initiation, development and ramification of infection threads within the nodule cortex or the division of infected plant cells. Infection thread development is closely linked to Nod factor substitutions (Ardourel et al., 1994; Walker & Downie, 2000) while nodule invasion is known to be mediated by bacterial surface polysaccharides (Niehaus et al., 1998; Pellock et al., 2000). Ineffective strains from this study may produce Nod factors that lack the correct substitutions to successfully colonize *P. vulgaris* nodules or may lack the right polysaccharides for enhanced nodule invasion.

Although further studies are clearly required to elucidate the determinants of effectiveness, the results indicate that the key limiting steps in *P. vulgaris* symbiosis are the infection and invasion steps. Analysis done ruled out bacteroid persistence and differences in N_2 fixation by bacteroids as key determinants of symbiotic outcomes in the *P. vulgaris*-rhizobia symbiosis.

3.4.3 PHB accumulation

Except CIAT 899, all strains accumulated large quantities of poly-β-hydroxybutyrate (PHB) in bacteroids, regardless of their effectiveness (Figure 3.7). PHB is a lipogenic product synthesized by rhizobia in growth-limiting conditions (Willis & Walker, 1998; Zevenhuizen, 1981) and has been associated with enhanced cell survival in the absence of an external carbon supply (Zevenhuizen, 1981), ineffectiveness (Sprent et al., 2013) and competitiveness (Willis & Walker, 1998). Mutants incapable of PHB production are either of enhanced effectiveness (Cevallos et al., 1996) or of similar effectiveness (Lodwig et al., 2005; Willis & Walker, 1998) to wild types.

In this study, no evidence was found linking PHB accumulation in bacteroids to effectiveness, which is contrary to suggestion by Sprent et al. (2013). PHB synthesis consumes carbon and reductant, but this carbon cost may be below the threshold detectable through differences in dry matter. Alternatively, the carbon cost of PHB may be compensated for by stimulation of N_2 fixation by the additional carbon sink (Kaschuk et al., 2009). From our observations, PHB accumulation is unlikely to be a major determinant of strain effectiveness in *P. vulgaris* symbioses.

3.4.4 *nodC* alleles

The *nodC* type harboured by a strain was found to be linked to its effectiveness. Strains carrying *nodC* type α and λ-b displayed higher rates of N_2 fixation in comparison to those with *nodC* types γ-a and γ-b (Figure 3.8). As already discussed in Chapter 2, NodC proteins are determinants of chain lengths of Nod factors produced by rhizobia in response to flavonoids. Since *nodC* genes are, in reported *Rhizobium* strains, located on the symbiotic plasmids, the genes may be considered as molecular markers of the general nature of symbiotic genes in a strain.

Previously, in extensive selection experiments done in Brazil under controlled conditions, the most effective *P. vulgaris* strains almost exclusively belonged to *R. tropici*

or affiliate species in the '*R. tropici* group' (Dall'Agnol et al., 2013; Hungria et al., 2003; Mostasso et al., 2002; Pinto et al., 2007). Strains in these species carry *nodC* type γ-b that was in the current study found to be associated with high effectiveness (Figure 2.5; Figure 3.8).

The other *nodC* type associated with high effectiveness was type α. The α allele of *nodC* is the dominant allele found in *R. etli* isolated from Mesoamerica (Aguilar et al., 2004), the origin of *P. vulgaris* (Bitocchi et al., 2012). In the current study, strains carrying *nodC* type α had the highest average rates of N_2 fixation (Figure 3.8). It is hypothesised that the efficiency of *nodC* type α results from the co-evolution history it may share with the host. In a study by Aguilar et al. (2004), Mesoamerican *P. vulgaris* nodulated preferentially, and sometimes even exclusively, with strains with *nodC* type α.

Representatives of the dominant group of strains nodulating *P. vulgaris* in Kenya, likely to belong to novel species, carried *nodC* type γ-a (and occasionally type γ-b) and were, in general, ineffective. As these strains are different from any described before, it is highly likely that these 'novel' groups acquired the nodulation genes through lateral transfer from other 'true' *P. vulgaris* symbionts. Their *nodC* sequences were 100% identical to those described in other rhizobia (Figure 2.5). In cases of recent lateral transfer of symbiotic genes, low rates of N_2 fixation in the recipients is not entirely surprising. In a study in Australia, resident *Mesorhizobium* strains that had acquired nodulation genes laterally from an inoculant strain were found to have lower rates of N_2 fixation (Nandasena et al., 2007).

The findings that *nodC* types α and γ-b were associated with higher effectiveness do not prove causation and, therefore, follow-up studies are required to establish causation. For example, nodulation genes, although transmissible, still tend to be tightly associated with the taxonomic positions of rhizobia. For example, in Chapter 2, *nodC* alleles were seen to be closely related to the taxonomic positions of the rhizobia. With notable exceptions, strains in the putative taxa mostly had *nodC* type γ-a, strains in species within the *R. etli* lineage had *nodC* type α and strains belonging to species in the *R. tropici* lineage had *nodC* type γ-b. With the link between nodulation genes and species, genetic differences between species may confound the interpretation of data obtained as genes unrelated to symbiotic genes may affect N_2 fixation outcomes. Examples of genes, related to N_2 fixation, carried in the chromosomes in members of the genus *Rhizobium*

include those that code for surface polysaccharides (Ormeño-Orrillo et al., 2012). The findings form a sound basis for interesting future studies.

3.4.5 Ability of BNF to meet N demand in *P. vulgaris*

Plants inoculated with CIAT 899, the most effective strain among those tested in this Chapter, accumulated approximately 40% the dry weight of the mineral N treatment that resulted in the highest SDW (Figure 3.11). The concentration of N in shoots of the plants inoculated with CIAT 899 was 4% (Figure 3.11), which was rather intriguing considering the N concentrations in shoots of plants inoculated with other effective strains (based on biomass assessment) were observed to be in the 3.7%-4.1% range (Figure 3.3).

When inoculated plants had N concentrations significantly higher than the range above, a positive linear relationship with dry matter ceased to exist (Figure 3.3). A similar trend was not observed in plants treated with mineral N, where and a direct relationship between N concentration and weight of plants continued to exist well into the 5%-6% N level (Figure 3.11).

One interpretation of the observed N concentration range in the inoculated plants is that N_2 fixation is regulated to keep N levels within an optimal range in plant tissues. Similar levels of N concentrations can be calculated for inoculated *P. vulgaris* in other studies. For example, shoot %N in inoculated *P. vulgaris* plants ranged between from 3.4% to 4.3% in the study by Buttery et al. (1997) and a maximum of 3.8% in Hungria and Kaschuk (2014). The optimal range of N shoot concentration may relate to a carbon: nitrogen ratio that regulates N_2 fixation (Hartwig, 1998). Considering that N concentrations in shoots correlate with the photosynthetic capacity of plants and therefore yield (Field & Mooney, 1986), autoregulation of this nature would place a cap on possible benefits from BNF in *P. vulgaris*.

Another important observation was that inoculated plants had reduced dry weights in comparison to those given mineral N, for a given N concentration in tissues (Figure 3.11), indicative of the energy cost of N_2 fixation in *P. vulgaris*. Nitrogen content, photosynthesis and CO_2 assimilation are linearly correlated (Evans, 1989; Field & Mooney, 1986). The observed differences in dry weights at the same shoot N concentration are likely from allocations of assimilated CO_2 to different sinks. In symbiotic plants, there is greater carbon sequestration towards nodulation (Minchin &

Witty, 2005). Also, N_2 fixation has a higher respiratory demand in comparison to nitrate reduction (Finke et al., 1982; Herridge et al., 2015; Minchin & Witty, 2005). From the data collected, it was clear that N_2 fixation is associated with lower N use efficiency -the ability of a genotype to acquire nutrients and use them for the production of biomass or utilizable material.

It is important to note that biomass accumulation is not perfectly correlated with grain yields as high supply of N stimulates biomass but can be limiting to N remobilization into pods and seeds (Baligar et al., 2001; Masclaux-Daubresse et al., 2010). Consequently, the results presented do not relate to seed yield, but only to biomass accumulation.

3.4.6 Scoring of nodulation and atypical nodules

A 4-point rating score for *P. vulgaris* was developed based on characteristics relevant to *P. vulgaris* nodulation (Appendix 4). Scores were found to correlate highly with shoot biomass (R^2=0.87, data not shown). Nodule rating scores are useful tools that provide data to corroborate other symbiosis measurements. Nodule rating systems have been developed for different legumes (Corbin et al., 1977; Yates et al., 2016a) but the utility of these rating scores across legumes is limited as diverse legumes exhibit different nodulation patterns.

NAK 120 was seen to induce a small fraction of oddly shaped nodules uncharacteristic of *P. vulgaris*. These nodules had smooth outer surfaces, lacked lenticels and were tapered. The oblong shape was in contrast to the spherical nodules usually seen on *P. vulgaris*. These nodules had a pink interior indicative of N_2 fixation (Figures 3.5E-F). While *Sesbania rostrata* has been reported to produce both determinate and indeterminate nodules (Fernández-López et al., 1998), a similar plasticity in nodule morphology has not been reported in *P. vulgaris*. NAK 120 was isolated from *Albizia sp.*, a legume tree in the tribe Ingeae (Brown 2008), a tribe that forms indeterminate ceasalpiniod nodules (Corby, 1988). NAK 120 belongs to *Rhizobium paranaense* (Chapter 2). It is speculated that the atypical nodules induced by NAK 120 result from incompatibilities between the strain and the host. The exact nature of the incompatibility remains unknown.

3.4.7 Concluding remarks

In addition to unearthing candidate inoculant strains, work in this chapter delivered significant contributions to the knowledge of the indicators of effectiveness in *P. vulgaris*-rhizobia symbiosis. PHB accumulation in bacteroids was found to be unrelated to strain effectiveness while the *nodC* genes, and the ability of a strain to colonise the nodule cortex, were found to be important determinants of symbiotic outcomes. These findings serve to narrow the research areas critical in the *P. vulgaris*-rhizobia symbiosis for further studies. Studies in this chapter also questioned the ability of biological N_2 fixation to meet the N demand of *P. vulgaris*. Evidence was found for an optimal shoot N concentration range in inoculated plants. With tissue N known to be closely linked to plant growth and yield, a regulation of N concentration in inoculated *P. vulgaris* to within the range observed would likewise place a cap on the maximum growth on yield benefits possible from inoculation. Data collected also hinted at low N use efficiency in inoculated plants.

In addition to inoculant strains being effective, they also need to be competitive. To conduct competition experiments, strain identification is required and the next chapter describes the development of a dual marker gene system for rhizobial competition studies.

CHAPTER 4
Using a broad-host-range plasmid containing *celB* and a *gusA* mini-transposon to assess rhizobial occupancy in *P. vulgaris* nodules

4.1 Introduction

The preclusion of an elite strain from some or all the nodules on a legume, by ineffective strain(s), leads to a reduction in the overall amount of N_2 fixed (Denton et al., 2002; Gerding et al., 2014; Thies et al., 1991a). Up to 59% of the yield variability in legumes following inoculation is reported to be related to nodule occupancy (Thies et al., 1991b), highlighting the need to screen potential inoculant strains for competitiveness and to understand rhizobial competition better.

Rhizobial competition studies require methods to reliably distinguish between rhizobial strains in the soil, on the rhizosphere and in the nodules. Standard approaches to distinguish strains rely on phenotypic, proteomic and genotypic differences among strains. Examples of these methods include melanin production, ELISA, MALDI-TOF, and RP01-PCR (Blanco et al., 2010; Gerding et al., 2014; Martínez-Molina et al., 2016; Spriggs & Dakora, 2009). The above methods suffer at least one of several drawbacks: insufficient discrimination, cross-reactions, expense, and/or labour intensiveness (Li et al., 2009; Ludwig, 2007). An alternative approach that sidesteps the need for intrinsic strain differences is the use of marker genes. Marker genes encode fluorescent proteins (Chalfie et al., 1994; Gage, 2002; Shaner et al., 2013) or enzymes that cleave chromogenic substrates into coloured products following oxidation (Sessitsch et al., 1998).

Two markers previously used in rhizobial competition studies are *gusA* and *celB* genes. The *gusA* gene encodes a β-glucuronidase (GUS) that cleaves glucuronides, in the presence of oxygen, into an insoluble coloured product (Jefferson et al., 1986). GUS activity is absent in rhizobia or from legumes, making *gusA* markers ideal for rhizobial nodule occupancy studies (Reeve et al., 1999; Sessitsch et al., 1996; Wilson et al., 1995). On the other hand, *celB* encodes a thermostable and thermoactive β-glucosidase with a high β-galactosidase activity (Voorhorst et al., 1995). Rhizobia and their hosts contain heat labile β-galactosidases, and a heat inactivation step is used to eliminate endogenous β-galactosidases before assay with β-galactosides (Sessitsch et al., 1996). The co-inoculation of plants with strains marked separately with the two markers allows the simultaneous detection of nodule occupants (Sanchez-Canizares & Palacios, 2013; Sessitsch et al., 1996).

The method of choice for introducing gene markers into rhizobia is through the use of mini-transposons (Anyango et al., 1998; Reeve et al., 1999; Sessitsch et al., 1996; Wilson et al., 1995). Transposons are efficient and easy to use, but their use necessitates that the mutants constructed are screened to ensure that competition for nodule occupancy or nodulation itself is not compromised by the transposon insertion (Sessitsch et al., 1998). Recently, Sanchez-Canizares and Palacios (2013) succeeded in separately marking two *Rhizobium* strains with *celB* and *gusA* at predetermined sites in the genomes, by fusion PCR and double recombination. This approach described above generated stable mutants but is still a laborious procedure that is not suited for use with a large number of strains and requires prior knowledge of genome of the target organism.

The delivery of marker genes into rhizobia can also be achieved through plasmids. Plasmids are easy to use but can be unstable in bacteria (Corich et al., 2001; De Gelder et al., 2007; Duodu et al., 2008; Gage et al., 1996) or they can generate an excessive metabolic load in the bacteria (Silva et al., 2011). However, stable broad-host-range plasmids have been used with rhizobia. One such example, pJP2, is stably maintained in *R. leguminosarum* (Karunakaran et al., 2005; Prell et al., 2002) *in planta* in the absence of selection. However, the efficiency of its *parCBA/DE* operon in the stabilization of plasmids in diverse rhizobia is unknown.

To facilitate the assessment of the competitiveness of rhizobia isolated from Kenyan soils for nodulation of *P. vulgaris*, I synthesised and cloned *celB* into the backbone of pJP2. The plasmid carrying *celB* was then conjugated into diverse rhizobia from Kenyan soils and investigated for stability. The marked strains were then assessed for competitiveness against CIAT 899, chromosomally marked with *gusA* using a mini-transposon.

This chapter reports an efficient dual-marker system, involving a plasmid and a mini-transposon, suitable for the marking of large numbers of rhizobia in the study of nodule occupancy in *P. vulgaris*. Further, the Chapter reveals the competitiveness of rhizobial strains from Kenyan soils against a leading inoculant strain.

4.2.1 Bacterial strains, plasmids and media

Strains and plasmids used in the study are listed and described in Table 4.1. *E. coli* strains were routinely cultured in lysogeny-broth (LB) media at 37°C and rhizobia in tryptone yeast (TY) media at 28°C (Hungria et al., 2016). For solid media, agar was added to 1.5% (w/v). All liquid cultures were incubated on a gyratory shaker set to 220 rpm. Media were supplemented with the following antibiotics and substrates as required (μg mL^{-1}): chloramphenicol (20), spectinomycin (100 for *E. coli* and 200 for rhizobia), streptomycin (100 for *E. coli* and 200 for rhizobia), nalidixic acid (75), tetracycline (20), X-Glc (50), X-Gal (50), and 5-aminolevulinic acid (50).

Table 4.1: Bacterial strains and plasmids used or constructed in this chapter

Strain/Plasmid	Relevant characteristics	Source/Reference
E. coli		
S17.1	Carrying pCAM131(mTn*5*SS*gusA*31) or pCAM111 (mTn*5*SS*gusA*11) λ*pir* Sm^R Sp^R Ap^R	Wilson et al. (1995)
ST18	*E. coli* S17.1 λ*pir* Δ*hemA* (ALA auxotroph)	Thoma and Schobert (2009)
Rhizobia		
R. tropici CIAT 899	Nx^R, Cm^R, Sp^S, Sm^R	Martinez-Romero et al. (1991)
R. leguminosarum bv. phaseoli 8002		Johnston et al. (1982)
NAK 091, NAK 103, NAK 104, NAK 120, NAK 157, NAK 210, NAK 214, NAK 220, NAK 227, NAK 231, NAK 239, NAK 242, NAK 245, NAK 254, NAK 266, NAK 284, NAK 287, NAK 288, NAK 295, NAK 299, NAK 303, NAK 312, NAK 315, NAK 321, NAK 327, NAK 332, NAK 334, NAK 351, NAK 358, NAK 363, NAK 367, NAK 368, NAK 378, NAK 382, NAK 387, NAK 407, NAK 458,	*Rhizobium* sp.	This work
Plasmids		
pJP2	*gusA,* Tc^R Ap^R	Prell et al. (2002)
pMK-RQ-CelBmNG	Km^R	This work
pGM01	Tc^R Ap^R, celBmNG	This work

4.2.2 Marking of CIAT 899 with *gusA*

4.2.2.1 Transfer of mini-transposons into CIAT 899

Two mini-transposons containing *gusA* - mTn*5*SS*gusA*11 (*gusA* driven by the constitutive *tac* promoter) and mTn*5*SS*gusA*31 (*gusA* driven by the symbiotically active *nifH* promoter)

were separately transferred into recipient CIAT 899 from the donor strain *E. coli* S17.1, an RP4 integrant, by bi-parental mating (Reeve et al., 2016). Briefly, CIAT 899 was cultured for 3 d to stationary phase in TY broth while *E. coli* S17.1, was started from a culture grown overnight with spectinomycin and streptomycin, and grown to log phase (~OD_{600nm} of 0.4) in LB broth with antibiotic selection. A 5 mL aliquot of each of the two cultures was separately concentrated by centrifugation (10,000 × *g* for 1 min), the pellet re-suspended in 200 μL saline (0.89% NaCl, w/v), mixed in a 1:1 ratio and a 100 μL aliquot spotted onto TY agar. The plate was incubated overnight at 28ºC and the culture spot re-suspended in 1 mL saline. Mutants were initially selected by plating the mating mix onto TY containing chloramphenicol, nalidixic acid, spectinomycin and streptomycin, before checking for GUS expression on TY with X-Glc agar. GUS-positive clones were stored in 15% (v/v) glycerol at 80ºC. Subsequently, antibiotic pressure was applied on mutants with spectinomycin and streptomycin.

4.2.2.2 Characterization of marked strains by morphology and growth rate

Two mutants from independent matings showing strong GUS activity on X-Glc plates were selected and named CIAT 899-G1 and CIAT 899-GE. CIAT 899-G1 carried mTn*5*SS*gusA*11 while CIAT 899-GE had mTn*5*SS*gusA*31. The two marked strains were streaked on TY plates (with and without antibiotic selection), and colony characteristics compared to those of the wild type.

To determine growth rate, a colony of wild-type CIAT 899, CIAT 899-G1 and CIAT 899-GE was inoculated separately into 5 mL TY broth (with antibiotic selection in the case of the two mutants) then incubated for 2 d. A 5 μL aliquot of the 2 d culture of each strain was then inoculated into 20 mL TY broth (with and without selection) in duplicates, and optical density at 600 nm assessed at nine intervals over a 27 h incubation period. At each point, OD_{600nm} values were read twice, and the mean value recorded before graphing of spectra-photometric values against incubation time. The mean generation time (MGT), in minutes, was calculated during log phase using the formula:

$$MGT = X \times \left(\frac{D}{Y-X}\right) \quad \text{(Eq. 4.1)}$$

Where X=OD_{600} at time 1, Y=OD_{600} at time 2 and D=duration in minutes between time 1 and time 2.

4.2.2.3 N_2 fixation and competitiveness of mutants

Wild-type CIAT 899, CIAT 899-GE and CIAT 899-G1 were assessed for capacity to fix N_2 on *P. vulgaris* cv. Kenya Tamu following procedures described in Section 3.2.2. The treatments were replicated four times, and plant shoots were excised 42 d after sowing, dried at 70ºC for 72 h then weighed. Roots were stained with X-Glc as per Section 4.2.2.4.

To assess competitiveness of marked strains for nodulation of *P. vulgaris* against the wild-type, all three strains were separately grown in 5 mL TY broth (with antibiotics for G1 and GE) to mid-log phase (OD_{600nm} of 0.5) then washed twice by pelleting (10,000 × *g*, 1 min) and re-suspending in deionised water. Subsequently, each of the cell pellets was re-suspended in 5 mL de-ionised water and OD_{600nm} adjusted with deionised water to give a final OD_{600nm} of 0.2. These cell suspensions were serial diluted to yield final concentrations of approximately 2×10^4 cells mL^{-1} followed by a 1:1 mixing of CIAT 899-G1 and CIAT 899-GE separately with the wild-type CIAT 899. A 1 mL aliquot of suspensions pre and post mixing was used for viable count determination on TY plates (with and without antibiotics) by the Miles and Misra drop plate count method (O'Hara et al., 2016). In the glasshouse, a 1 mL aliquot of the mixtures was applied per pre-germinated seed. Three seeds were inoculated and sown per pot but thinned to two plants after 7 d. Each treatment, including the un-inoculated, had three pot replicates. Other procedures and maintenance of plants in the glasshouse were as described in Section 3.2.2. Plants were harvested after 21 d, roots stained with X-Glc (4.2.2.4), followed by scoring of nodules by colour. The number of blue or unstained nodules was counted and a competitive index (CI) calculated as, the mutant-to-wildtype ratio in the nodules divided by the corresponding ratio in the inoculum.

$$\text{CI of Y}=\left(\frac{NodulesY}{NodulesX}\right) / \left(\frac{cfuY}{cfuX}\right) \qquad \text{(Eq. 4.2)}$$

CI is competitiveness index; Y is the mutant strain, its nodules or colonies; X is the wildtype, its nodules or colonies; and cfu is colony forming units mL^{-1} of inoculum.

4.2.2.4 Nodule staining procedure

Staining of roots for GUS was performed as described by Wilson et al. (1995). Each root was vacuum-infiltrated for 30 min with 200 mL GUS staining buffer (50 mM $NaPO_4$, 1 mM EDTA (pH 8.0), 0.1% (v/v) Triton X-100, 0.05% (w/v) SDS, 0.1% (w/v) Sarkosyl and 200 μg mL^{-1} of X-Glc) prior to a 24 h incubation at 37°C with agitation (200 rpm). Stained roots were then destained in 4% (w/v) sodium hypochlorite for 30 min, rinsed and stained nodules counted.

4.2.2.5 Determination of mTn*5*SS*gusA*31 insertion-site in CIAT 899-GE

Chromosomal DNA flanking mTn*5*SS*gusA*31 was identified by a method modified from the semi-random two-step PCR protocol of Chun et al. (1997). The modified method involved three rounds of PCR (Figure 4.1). Genomic DNA was extracted from CIAT 899-GE using Promega's Wizard® DNA Purification Kit according to manufacturer's instructions and 3 μL (~120 ng) used as template for PCR 1. PCR 1 was performed using GUS 134 primer (Table 4.2) in a 25 μL reaction (Table 4.3) using the PCR 1 program (Table 4.4). The product of PCR 1 was diluted 5-fold with PCR grade water and 1 μL used as the template for PCR 2. PCR 2 was carried out in a 25 μL reaction (Table 4.3) with GUS 134 primer paired with, separately, CEKGRNB1, CEKGRNB2, CEKGRNB3 or CEKGRNB4 (Table 4.2). Reactions were carried out using PCR 2 program (Table 4.4). 1 μL aliquots of 5-fold dilutions of PCR 2 products were used as DNA templates for PCR 3 performed in 25 μL reactions (Table 4.3) with WIL3 primer and CEKG4 primer (Table 4.2) following PCR 3 program (Table 4.4).

Products of PCR 3 were separated in 1% (w/v) agarose, excised and purified, as described in Section 2.2.3.3. Purified amplicons were sequenced by Sanger technique using WIL3 and CEKG4 primers through the Australian Genome Research Facility (AGRF).

Homology searches and sequence comparisons were performed with the sequence data obtained using the NCBI BLASTN search algorithm to identify transposon sequences and CIAT 899 sequences. Primers were designed, to confirm the mTn*5* insertion site, to span either flanking DNA and IE or flanking DNA and OE (Table 4.5). The primers were used in 25 μL PCR reactions containing: 12.5 μL of 2 × GoTaq® Green Master Mix [(pH 8.5), 400 μM dATP, 400 μM dGTP, 400 μM dCTP, 400 μM dTTP and 3 mM $MgCl_2$], 0.4 μM IE-R/OE-R primer, 0.4 μM IE-F/OE-F primer, 80 ng genomic DNA and UltraPure PCR grade water (Fisher Biotec) to 25 μL. PCR thermal cycling

conditions were: initial denaturation at 95ºC for 120 s; 32 cycles of 94°C for 45 s, 55°C for 60 s and 72°C for 90 s; and a final extension at 72°C for 5 min. Products were separated on 1% agarose gel, purified and sequenced with IE-F and OE-F primers through the Australian Genome Research Facility (Section 2.2.3.3). BLASTN was used to search for highly similar sequences in GenBank to those obtained.

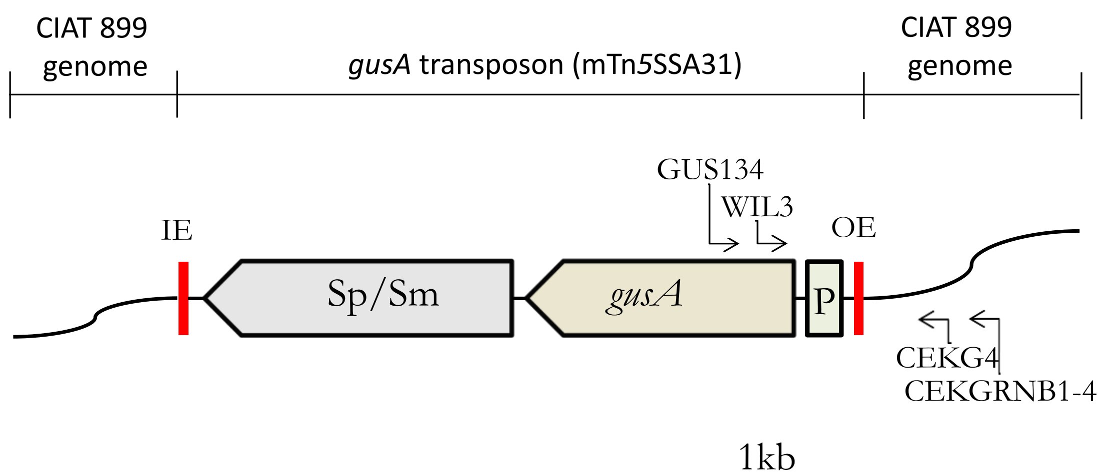


Figure 4.1: Primer binding sites in PCR strategy to determine the insertion point of mTn5SS*gusA*31 in CIAT 899-GE genome. Primers are not drawn to scale. P-promoter, IE-inner end, OE-outer end, Sp/Sm-genes for spectinomycin and streptomycin resistance, *gusA*-promoterless *gusA* coupled to a promoter (P).

Table 4.2: Sequences of oligonucleotides used

Primer	Sequence	Source/Reference
CEKG4	5'-GGCCACGCGTCGACTAGTAC-3'	Chun et al. (1997)
CEKGRNB1	5'-GGCCACGCGTCGACTAGTACNNNNNNNNNNGCGCGC-3'	W. Reeve#
CEKGRNB2*	5'-GGCCACGCGTCGACTAGTACNNNNNNNNNNGCCGCC-3'	This study
CEKGRNB3*	5'-GGCCACGCGTCGACTAGTACNNNNNNNNNYICCGCC-3'	This study
CEKGRNB4*	5'-GGCCACGCGTCGACTAGTACNNNNNNBBBBNCCGCC-3'	This study
GUS 134	5'-CTT GTA ACG CGC TTT CCC AC-3'	W. Reeve#
WIL3	5'-GAATGCCCACAGGCCGTCGAG-3'	Wilson et al. (1995)

ACGT-Standard nucleotides; N-any; Y-C or T; B-C or G or T; I-Inosine

*The specific sequences on the 3' end that follow the degenerate were generated by searching for motifs approximately every 1000 bp in the genome of CIAT 899 (GenBank: CP004015.1) and provide an initial clamp near the transposable element to allow initial amplification.

#Murdoch University, Perth, WA

Table 4.3: PCR reaction components for a 25 μL reaction

PCR 1		PCR 2		PCR 3	
PCR reagent	**μL**	**PCR reagent**	**μL**	**PCR reagent**	**μL**
PCR water	9	PCR water	10.5	PCR water	10.5
2× GoTaq Green	12.5	2× GoTaq Green	12.5	2× GoTaq Green	12.5
GUS 134 (50 μM)	0.5	GUS 134 (50 μM)	0.5	WIL3 (50 μM)	0.5
DNA (40 ng μL^{-1})	3	CEKGRNB1-4 (50 μM)	0.5	CEKG4 (50 μM)	0.5
		5-fold dil PCR1	1	5-fold dil PCR2	1

Table 4.4: Thermal cycling conditions for typing of transposon insertion site

PCR 1		PCR 2		PCR 3	
Conditions	**Cycles**	**Conditions**	**Cycles**	**Conditions**	**Cycles**
94°C 2 min	×1				
94°C 30s 60°C 30s 70°C 90s	×10	94°C 30s 42°C 30s 70°C 3 min	×1	94°C 30s 60°C 30s 70°C 90s	×30
		94°C 30s 41°C 30s 70°C 3 min	×1	70°C 7 min	×1
		94°C 30s 40°C 30s 70°C 3 min	×1		
		94°C 30s 39°C 30s 70°C 3 min	×1		
		94°C 30s 38°C 30s 70°C 3 min	×1		
		94°C 30s 60°C 30s 70°C 3 min	×24		

Table 4.5: Sequences of oligonucleotides used to amplify regions flanking OE and IE in CIAT 899-GE

Primer	Sequence	Target region	Reference
IE-F	5'-CGATTGCCTTGAACTCACGG-3'	Xanthine phosphoribosyltransferase	This study
IE-R	5'-CGAAGTAATCGCAACATCCGC-3'	Sp/Sm region of mTn*5*SS*gusA*31	This study
OE-F*	5'-CTTGTAACGCGCTTTCCCAC-3'	*gusA* of mTn*5*SS*gusA*31	W. Reeve#
OE-R	5'-GAACGGCTCCAAGGAAGTGG-3'	Universal Stress Protein A	This study

*Same as GUS134

#Murdoch University, Perth, WA

4.2.3 Marking of rhizobia with *celB* gene

4.2.3.1 Construction of a stable, broad-host-range CelB plasmid

DNA sequences coding for CelB (Voorhorst et al., 1995) and mNeonGreen (Shaner et al., 2013) proteins were obtained from GenBank and edited in Geneious® software to remove unwanted restriction sites and incorporate XbaI and PstI sites on the flanks for cloning. CelB and mNeonGreen amino acid sequences were maintained and put under the control of a constitutive *tac* promoter. Once *in silico* design was complete, the sequence was sent to GeneArt™ (ThermoFisher Scientific) where the synthetic genes were assembled from synthetic oligonucleotides. The 2,369 bp fragment was cloned into pMK-RQ (Km^R) to create pMK-RQ_celBmNG (Figure 4.2) and final construct verified by sequencing. pMK-RQ_celBmNG was digested with PstI and XbaI, releasing a 2,337 bp fragment that was gel purified and ligated into pJP2 digested with PstI+XbaI, creating pGM01.

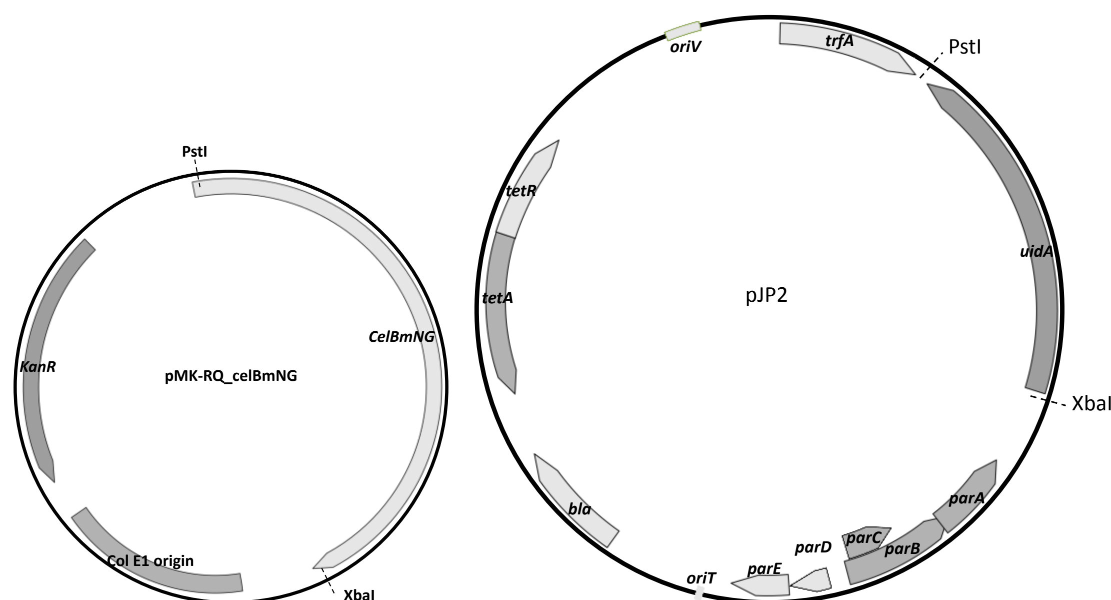


Figure 4.2: Maps of pMK-RQ_celBmNG and pJP2. In pJP2, the promoter-probe vector encodes genes for tetracycline and ampicillin resistance. In addition to these, RK2 genes *trfA*, *oriV* and *oriT* are incorporated as well as *par* genes for plasmid stability (Prell et al., 2002). Note the PstI and XbaI sites that enabled the excision of CelBmNG region from pMK-RQ_celBmNG into the *uidA* region of pJP2.

Restriction digestions and ligation mentioned above were carried out largely according to instructions by the manufacturer of enzymes (Promega). Briefly, double restriction enzyme digests with XbaI and PstI of pMK-RQ_celB-mNG and pJP2 were carried out in 20 μL reactions containing the following at indicated concentrations: 1 × of Promega's restriction buffer H, 2 μg μL^{-1} acetylated BSA, 10ng μL^{-1} DNA and nuclease free water to make up to 19 μL. A 0.5 μL aliquot of each of the two enzymes was then added followed by digestions for 4 h at 37°C. Enzymes were heat inactivated at 74°C for 15 min prior to separation of fragments in 1% (w/v) agarose gel, excision and purification (Section 2.2.3.3). For ligation, a 20 μL, 1:2 vector to insert, ligation mix contained the following: 2 μL (c. 100ng) pJP2 (with *uidA* removed with PstI and XbaI), 10 μL (~200 ng) of insert, 2 μL T4 ligase buffer, 5.67 μL water and 0.33 μL T4 DNA ligase. Ligation was carried out overnight at room temperature (approx. 15°C) and products (pGM01) transformed into *E. coli* ST18.

4.2.3.2 Transformation of pGM01 into *E. coli* ST18

Before transformation, electrocompetent ST18 cells were prepared as follows. ST18 was cultured overnight in 50 mL LB with 5-aminolevulinic acid (ALA) at 28ºC with shaking (250 rpm). The overnight culture was sub-cultured into four 250 mL LB-ALA flasks and incubated to an OD_{600nm} of 0.2. Cultures were placed on ice, pelleted at 4°C by centrifuging (1,000 × *g*) for 20 min, the supernatant decanted and the pellet re-suspended in 15 mL of sterile 10% (v/v) glycerol with 20 mM β-mercaptoethanol. This process was repeated twice with re-suspension in 15 mL of 10% (v/v) glycerol. Finally, the pellet was re-suspended in 200 μL of 10% (v/v) glycerol, and aliquots of 40 μL transferred into 1.5 mL tubes, frozen in dry ice and stored at -80C until required.

The plasmid pGM01 was transformed into ST18 by electroporation. A 40 μL aliquot of competent cells was thawed on ice and mixed with 10 ng DNA from ligations above. The mixture was transferred into a chilled 0.1 cm cuvette and sample electroporated using a Biorad Genepulser® II electroporator (2.5 kV, 200 Ω, 25 μF) for 5 s. To the cells in the cuvette, 1 mL SOC medium (0.5% (w/v) yeast extract, 2% (w/v) tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgSO_4$ and 20 mM glucose) supplemented with ALA was added and solution transferred into a 15 mL snap-cap tube before 2 h of incubation at 37°C with shaking (200 rpm). Aliquots of 10 μL, 100 μL and 200 μL were spread-plated onto LB plates with selection (50 μg μL^{-1} tetracycline) and X-Gal (50 μg

μL^{-1}). Colonies appearing on selection plates were screened for pGM01 presence as detailed in 4.2.3.4 by PCR and by a β-galactosidase assay.

4.2.3.3 Conjugal transfer of pGM01 into rhizobia

The plasmid pGM01 was transferred from ST18 into 39 rhizobial strains (Table 4.1) by bi-parental mating as per Section 4.2.2.1. Before mating, ST18 carrying pGM01 was cultured in LB supplemented with tetracycline and ALA. Mating mixes were spotted on TY plates with ALA for overnight incubation and selection of transconjugants was on TY supplemented with tetracycline. All other steps were as described in Section 4.2.2.1. For the designation of transconjugants carrying pGM01, a capital letter C was added to strain codes.

4.2.3.4 Confirmation of presence of pGM01 and expression of thermostable CelB

In addition to growth on tetracycline plates, the presence of pGM01 in ST18C and rhizobia was confirmed by PCR. Primers PGM949F (5'-GGAGAAGTACCGC AAGCTGT-3') and PGM1537R (5'-CCGTTCTCTGGTAGATCGCC-3') were designed to amplify a 589 bp region internal of the *celB* gene on pGM01. A 25 μL PCR reaction was carried out with the following components in the stated final concentration: GoTaq® Green Master Mix (1×); PGM949F (1.0 μM); PGM1537R (1.0 μM); DNA template (one colony) and nuclease free water to 25 μL. Thermal cycling conditions for the PCR were: initial denaturation 94°C for 4 min; 30 cycles of 94°C for 45 s, 55°C for 45 s and 72°C for 60 s; and a final extension at 72°C for 7 min. PCR products were separated in 1% (w/v) agarose gel (Section 2.2.3.3.)

To confirm CelB expression and activity, a β-galactosidase assay modified from that described by Terpolilli (2009) was carried out. Strains carrying pGM01 were grown in TY or LB broth with selection. Unmarked rhizobia and *E. coli* BW 20767, acting as controls, were cultured without antibiotic selection. At log phase, 1.5 mL aliquots of the cultures were separately pelleted (10,000 × *g*, 1 min) followed by suspension of cells in 1 mL of enzyme buffer [50 mM $NaPO_4$ (pH 7.0), 1 mM EDTA (pH 8.0)]. The suspensions were pipetted in equal halves into 1.5 mL tubes, one-half of the tubes incubated at 70°C for 45 min while the remaining half was left to sit on the bench at room temperature. 50 μL chloroform and 25 μL SDS (0.1%, w/v) were then added to all tubes and tubes vortexed for 10 s. A 50 μL aliquot was removed from each tube and separately added to 450 μL buffer Z (50 mM $NaPO_4$, 1 mM EDTA, 10 mM

mercaptoethanol, 10 mM ONPG) and incubated for 30 min at 37°C. At the end of the incubation period, 500 μL of 1 M Na_2CO_3 was added to stop the reaction and presence or absence of yellow colour noted or determined at OD_{420nm}. An alternative method used was to re-suspend duplicate cultures in enzyme buffer, incubate half at 70°C for 45 min, add chloroform and SDS to lyse as above and finally add Magenta-Gal and incubate at 37°C for 30 min.

4.2.3.5 Characterization of transconjugants by morphology and growth rate

The 39 rhizobial strains marked with pGM01 and their unmarked parents were streaked on TY with and without selection and after 4 d of incubation, compared in their colony characteristics. Additionally, a subset of seven strains was selected from the 39 for comparison of growth rates between parental strains and pGM01 transconjugants. Representative strains selected were NAK 120, NAK 103, NAK 334, NAK 210, NAK 287, NAK 239 and CIAT 899. Growth curves and growth rates were determined in antibiotic free TY broth as described in Section 4.2.2.2 by sampling 14 times for spectrophotometric measurements over a 27h period.

4.2.3.6 Plasmid stability in the absence of selection

Three strains carrying pGM01 were selected to test the ability of pGM01 to be maintained in rhizobia in the absence of antibiotic selection. Sub-cultured from stationary phase cultures, NAK 334C, NAK 120C and NAK 210C were grown into log phase (OD_{600nm} of 0.5) in TY broth with tetracycline. From these cultures, 5 μL aliquots were separately and in duplicates sub-cultured into 20 mL TY broth without antibiotics, mixed and 1 mL aliquot immediately removed before incubation of remainder at 28°C on a shaker set at 220 rpm. The 1 mL aliquot was serially diluted for determination of CFU mL^{-1} on TY (without tetracycline) by the Miles and Misra plate count method (O'Hara et al., 2016). One dilution from the Miles and Misra process was spread on TY, incubated and 100 isolated colonies replica patched on TY and TY-tetracycline plates. Growth on TY-tetracycline plates indicated the presence of pGM01.

Incubated cultures were grown to log phase (OD_{600nm} of 0.5), and two aliquots removed. The first, a 1 mL aliquot was used for viable cell count by the Miles and Misra plate count, spread plating and replica patching as described above. The second, a 5 μL aliquot was sub-cultured and mixed into fresh 20 mL TY broth before the immediate removal of a 1 mL aliquot for a viable count on TY, spread plating and replica patching

as earlier described. This process of culturing and sub-culturing while counting total populations and tetracycline resistant proportions, at start and end times was repeated eight times.

Comparison of the number of cells present at the start of culture and end of culture allowed determination of the number of generations elapsed for each culture period using the formula.

$$\text{Generations elapsed} = \log(CFU\ per\ mL\ at\ end) - \log(CFU\ per\ mL\ at\ start)\ /log2 \quad \text{(Eq. 4.3)}$$

4.2.4 Competitiveness of *celB*-marked strains against CIAT 899-GE for nodulation of *P. vulgaris*

All strains marked with *celB* were separately cultured in 5 mL TY broth with tetracycline until log phase (OD_{600nm} of 0.5). CIAT 899-GE (*gusA*) was similarly cultured in TY containing spectinomycin and streptomycin while wild-type CIAT 899 was grown in antibiotic-free TY broth. A 1 mL aliquot of each culture was pelleted, washed twice free of media with sterile deionised water by centrifugation (10,000 × *g*, 1 min) and cells re-suspended in 1 mL sterile deionised water. The optical density of suspensions was determined and using sterile deionised water, volumes adjusted to give an OD_{600nm} of 0.1 (theoretical 1×10^8 cells mL^{-1}) before a serial dilution to give a theoretical concentration of approximately 1×10^4 cells mL^{-1}. All strains at this concentration were mixed 1:1 with a CIAT 899-GE suspension of a similar concentration and 1 mL aliquots of mixtures separately applied on pre-germinated seeds. A 1 mL aliquot of each mixture was spared for viable cell counts on antibiotic free TY agar and TY with appropriate antibiotics by the Miles and Misra plate count. Two seeds were inoculated and sown per pot filled with steam-sterilized sand and treatments were replicated three times. Plants were maintained in the glasshouse as described in Section 3.2.2, harvested 21 d after inoculation and roots stained.

Roots were first stained with GUS buffer (Section 4.2.2.4) and following overnight incubation at 37°C, incubated in a 70°C oven for 2 h to destroy endogenous β-galactosidases. CelB staining was then done by vacuum-infiltrating each root for 30 min with 200 mL Magenta-gal staining buffer (50mM $NaPO_4$, 1mM EDTA (pH 8.0), 0.1% (v/v) Triton X-100, 0.05% (w/v) SDS, 0.1% (w/v) Sarkosyl and 200 μg mL^{-1} Magenta-gal) prior to a further 24-hour incubation at 37°C with agitation (200 rpm). Stained roots were then de-stained as earlier described (4.2.3.4) and scored for colour.

Competitiveness index of strains for nodulation of *P. vulgaris* was calculated using the formula in Section 4.2.2.3.

4.2.5 Data analysis

Where applicable, data were subjected to an analysis of variance (ANOVA) using SPSS version 22 (IBM Corp, released 2013). ANOVA was preceded by a test for normality and equal variances (Levene's test). Least significant difference (LSD) was then used when ANOVA was found to be significant.

4.3 Results

4.3.1 Marking of *R. tropici* CIAT 899 with *gusA*

The two mini-transposons carrying *gusA* genes were readily mobilized from the donor strain *E. coli* S17.1 into CIAT 899, as evidenced by the growth and isolation of colonies on TY plates supplemented with nalidixic acid, chloramphenicol, spectinomycin and streptomycin. On X-Glc plates, the CIAT 899 mutants also expressed GUS, a phenotype not observed in wild-type CIAT 899. Colonies of mutants marked with *gusA* under the constitutive P*tac* (mTn*5*SS*gusA*11), developed a blue hue on X-Glc plates while those carrying a symbiotically active *gusA* driven by P*nifH* (mTn*5*SS*gusA*31) required several days of incubation to turn blue (Figure 4.3).

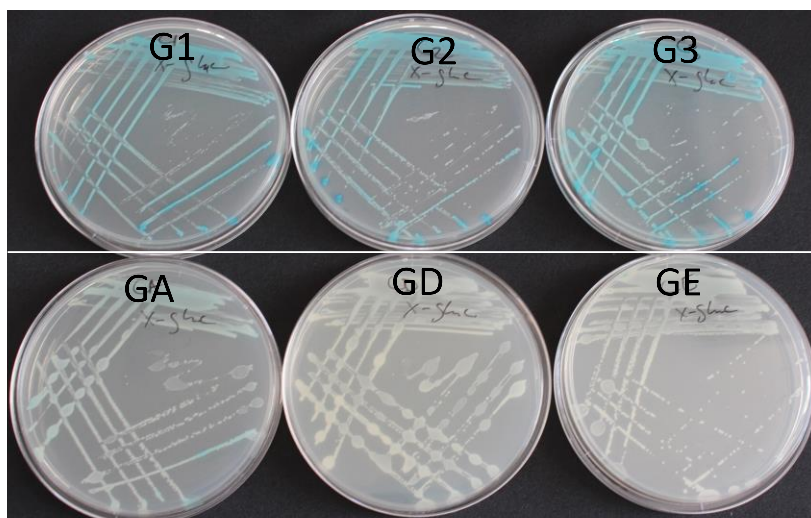


Figure 4.3: Differential staining of CIAT 899 marked with *gusA* growing on TY media supplemented with X-Glc after 2 d of incubation. G1, G2, G3 carry a constitutively expressed *gusA* gene while GA, GD, GE carry a symbiotically active *gusA* gene. Colonies of GA, GD and GE, required several days of incubation to turn blue.

4.3.1.1. Effect of mini-transposons on growth and morphology of marked strains

Two GUS mutants of CIAT 899 namely, CIAT 899-C1 and CIAT 899-GE, were selected from the many mutants obtained and assessed for similarities, in colonial morphologies, to wild-type CIAT 899 on plain TY. Following 4 days of incubation at 28ºC, colonies of all three strains were found to be 1.5 to 2.5 mm in diameter, circular with smooth margins, semi-translucent, elevated and very mildly mucoid (Figure 4.4). No differences were noted in the characteristics of the colonies.

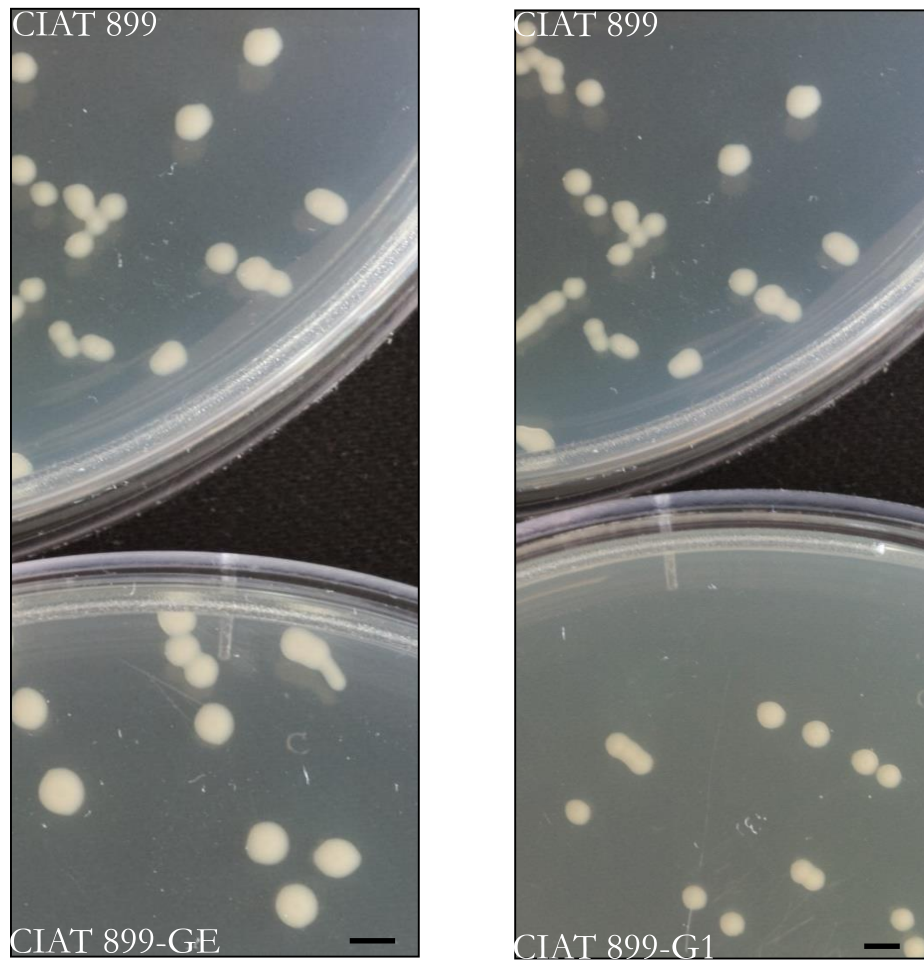


Figure 4.4: Colonies of *R. tropici* CIAT 899 and its two *gusA* mutants G1 and GE after 4 d on TY agar plates at 28ºC. Scale bar is 2 mm.

In antibiotic-free TY broth media, CIAT 899, CIAT 899-GE and CIAT 899-G1 had growth curves typical of batch cultures (Figure 4.5). The three strains experienced a lag phase that lasted between 5 to 10 h with CIAT 899-GE displaying a relatively a longer lag period. CIAT 899-GE also reached a maximum optical density at a relatively lower value than the wild-type and CIAT 899-G1. However, the slopes during log phase were quite similar and mean generation times calculated from the curves were 111 min for wild type CIAT 899, 119 min for CIAT 899-G1 and 112 min for CIAT 899-GE, and were not significantly different from each other ($p>0.05$).

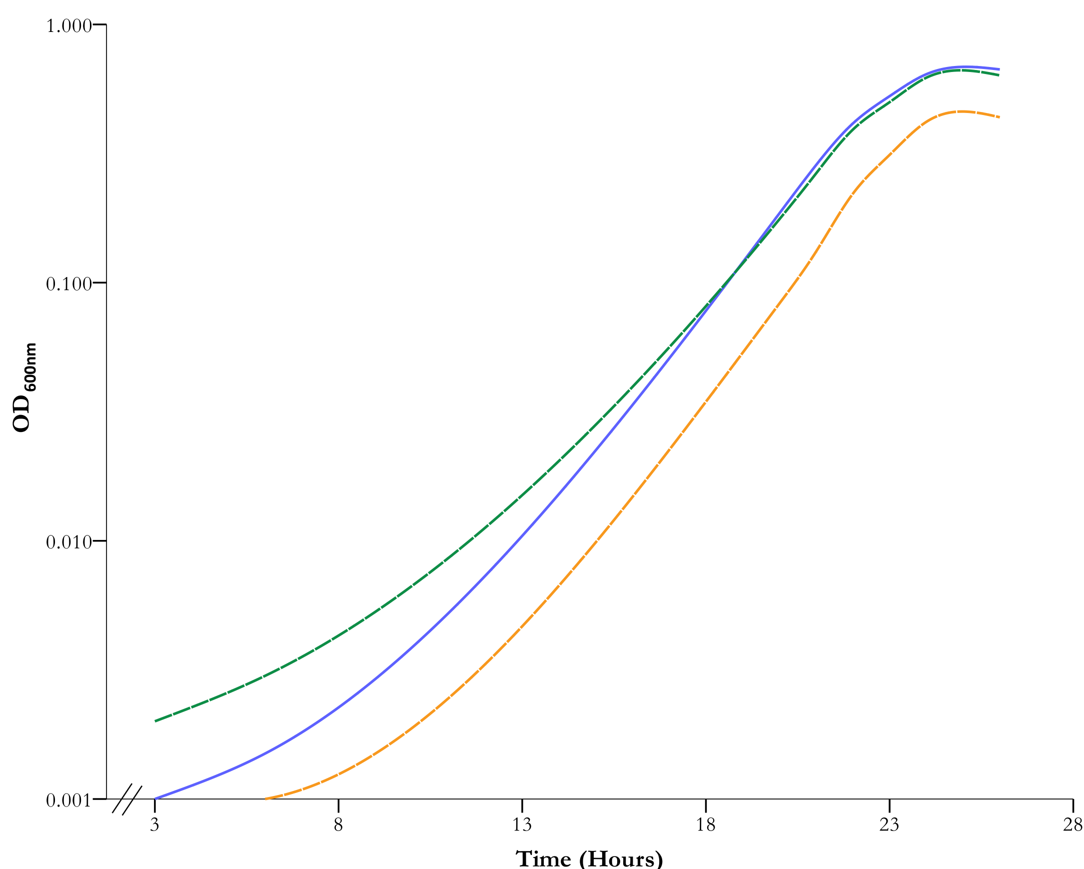


Figure 4.5: Representative growth curves (OD_{600nm} versus time in hours) of wild-type CIAT 899 (—) and two of its *gusA* mutants, CIAT 899-G1 (—-—) and CIAT 899-GE (—-—). OD_{600nm} values of cultures were taken at nine intervals over 27 h and used to plot the graphs. Mean generation times, in minutes, calculated during log phase were: CIAT 899=111, CIAT 899-G1=119 and CIAT 899-GE=112. The mean generation times of the mutants did not differ significantly from that of the wild-type ($p>0.05$).

4.3.1.2 Effect of mini-transposons on N_2 fixation and competitiveness of marked strains

CIAT 899-G1, CIAT 899-GE and wild-type CIAT 899 induced pink nodules on *P. vulgaris* cv. Kenya Tamu. Plants in the un-inoculated treatment did not form any nodules. On staining with X-Gluc, nodules resulting from inoculation with CIAT 899-G1 and CIAT 899-GE were blue, while those from inoculation with the wild type remained unstained. In addition to nodules, background staining of roots was also observed in plants inoculated with the CIAT 899-G1 (Figure 4.6). In the CIAT 899-GE treatment in which no root staining occurred (Figure 4.6). CIAT 899-G1 has *gusA* under

the constitutive P*tac* promoter while CIAT 899-GE carries *gusA* under P*nifH*, which is symbiotically active.

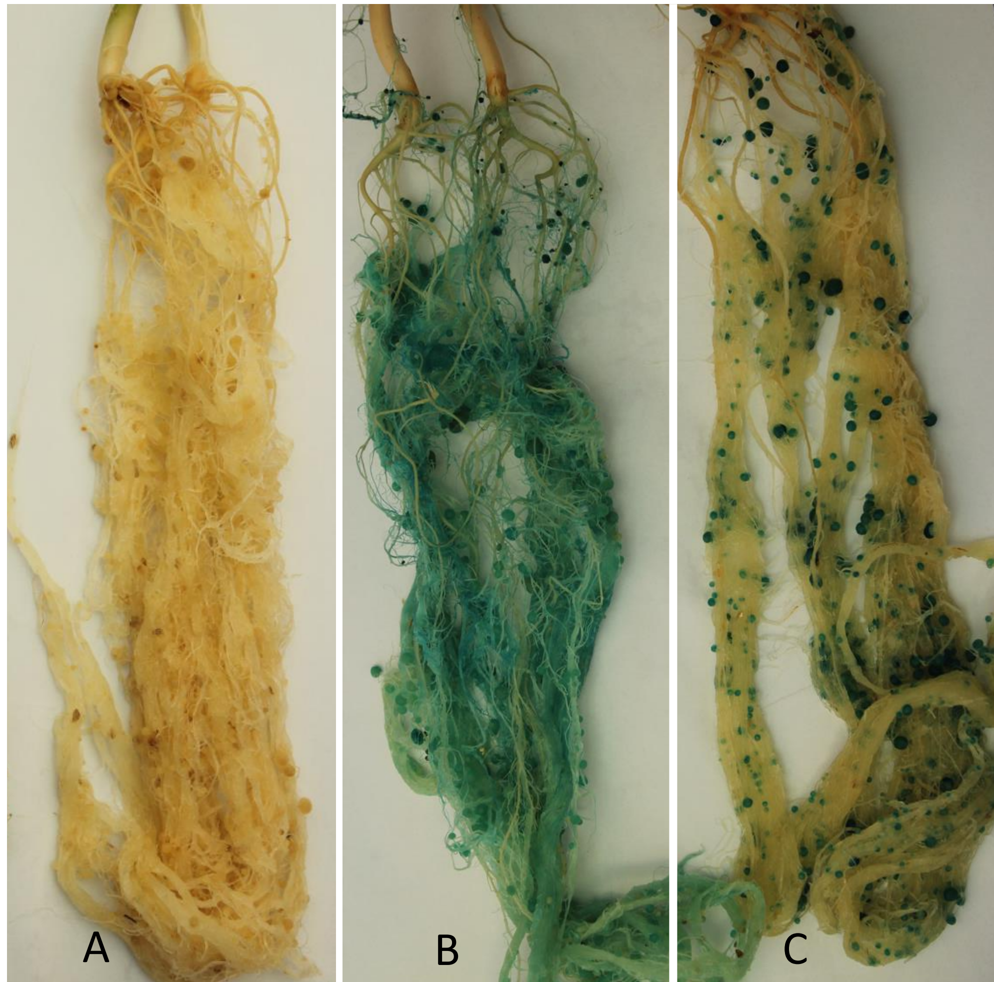


Figure 4.6: Roots of *P. vulgaris* cv. Kenya Tamu inoculated with, **A**, wild-type CIAT 899, **B**, CIAT 899-G1 and **C**, CIAT 899-GE after 24-h staining in X-Glc solution. In addition to nodule staining, plants inoculated with CIAT 899-G1 showed high levels of background root staining (A). Plants inoculated with CIAT 899-GE (B) showed only nodule staining while those inoculated with the wild-type had unstained roots and nodules (C).

The mean shoot dry weight of a plant inoculated with CIAT 899-G1 was 3.61 ±0.20 g, 3.54 ±0.39 g for CIAT 899-GE and 3.28 ±0.23 g for the wild type. The differences in the three means were not statistically significant (LSD, $P>0.05$), indicating that the insertions did not compromise the N_2-fixing abilities of the two marked strains.

In the experiment to test the competitiveness of CIAT 899-G1 and CIAT 899-GE against wild-type CIAT 899 for nodulation of *P. vulgaris*, the CIAT 899-G1 + CIAT 899 inoculum applied per seed (~1.4 $\times 10^4$ cells) at sowing contained 49.7% G1 and 50.3% wild type cells. The CIAT 899-GE + CIAT 899 inoculum applied per seed (~1.6 $\times 10^4$) had 50.2% of GE and 49.8% of wild-type cells. Upon harvest, roots contained both blue and unstained nodules following incubation in X-Glc. Blue nodules were scored as being occupied by the *gusA*-marked CIAT 899-G1 or CIAT 899-GE while unstained nodules were scored as being occupied by the wild-type CIAT 899. Less than 10% of nodules had both blue and unstained portions, indicative of dual occupation, and these were omitted from consideration when analyzing competitiveness. Plants in the un-inoculated treatment did not form any nodules. Mean proportions of nodules occupied by the marked strains were 52.3% in the CIAT 899-G1 + CIAT 899 treatment and 50.3% in the CIAT 899-GE + CIAT 899 treatment. CIAT 899-G1 and CIAT 899-GE were found to be of equal competitiveness to wild type CIAT 899 for nodulation of *P. vulgaris* 'KT'. The two marked strains were also of equal competitiveness ($P>0.05$).

4.3.1.3 Insertion site of mTn*5*SS*gusA*31 in the genome of CIAT 899-GE

The location of the mini-transposon (mTn*5*SS*gusA*31) in the genome of CIAT 899-GE was identified by a semi-random PCR strategy. The second PCR of the 3-PCR protocol yielded similarly sized products (Figure 4.7), but re-amplification of these with internal primers in PCR 3 yielded differently sized products (Figure 4.7).

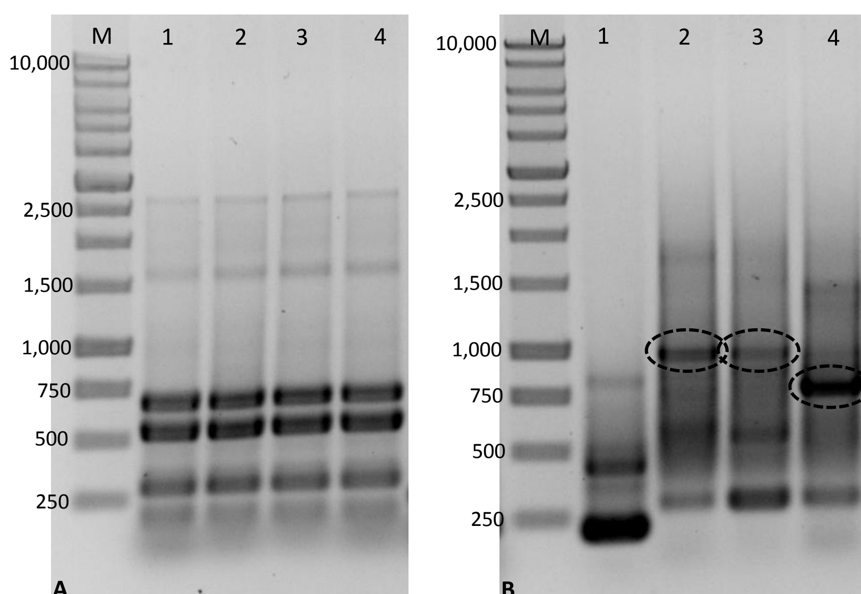


Figure 4.7: Images (on previous page) of agarose gel electrophoresis of products of PCR 2 (**A**) and PCR 3 (**B**) in the typing of the mTn*5*SS*gusA*31 insertion site in the CIAT 899-GE genome. In **A**, Lanes 1 to 4 are reactions with GUS 134 primer paired with CEKGRNB1, CEKGRNB2, CEKGRNB3 and CEKGRNB4, respectively. In **B**, Lanes 1 to 4 are reactions of products in corresponding lanes in image A with WIL3 and CEKG4 primers. Circled bands were excised for sequencing. M=1 kb DNA ladder (Promega)

Fragments in the expected size range (700 bp-1,000 bp) were excised from the gel for sequencing and alignments in Geneious® revealed a 904 bp consensus sequence. This region was analyzed by BLASTN (discontiguous megablast). A 64 bp region aligned with *gusA* sequences, 257 bp with *nifH* sequences and 556 bp with the universal stress protein gene (*uspA*) of CIAT 899 (Figure 4.8).

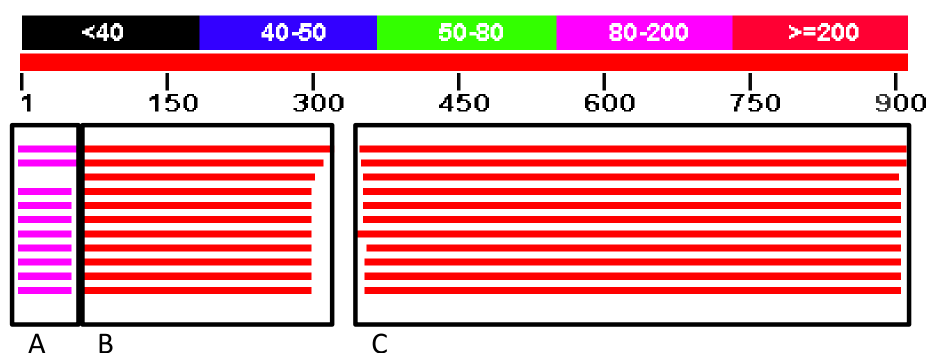


Figure 4.8: BLASTN results, showing the distribution of blast hits on the 904 bp consensus query sequence. Bases 1-64 (A) aligned with *gusA* sequences, 65-321 (B) aligned with *nifH* sequences of *Rhizobium* and *Bradyrhizobium* while 349-904 aligned with *uspA* in *Rhizobium* and *Bradyrhizobium*. The 27 bp region between B and C did not find matches. The 27 bp included the 19 bp repeat flanking mTn*5*SS*gusA*31 on the OE.

The BLAST results suggested the mini-transposon had inserted 6 bp from the 3' end of *uspA* gene, splitting the 846 bp gene into two parts of 6 bp and 840 bp. To confirm this, insertion junctions at both the IE and the OE ends of mTn*5*SS*gusA*31 were amplified using primers designed from the CIAT 899 genome (GenBank: CP004015.1). Primer IE-F (Table 4.5), complementary to Xanthine phosphoribosyltransferase gene downstream of *uspA* in the CIAT 899 genome, and IE-R (Table 4.5), complimentary to Sp/Sm region of mTn*5*SS*gusA*31, were successfully used to amplify 750 bases around the IE (Figure 4.9). Similarly, primer OE-F (Table 4.5), binding to *gusA* in mTn*5*SS*gusA*31, and OE-R

(Table 4.5), binding to *uspA*, amplified 800 bp around the OE junction in CIAT 899-GE (Figure 4.9).

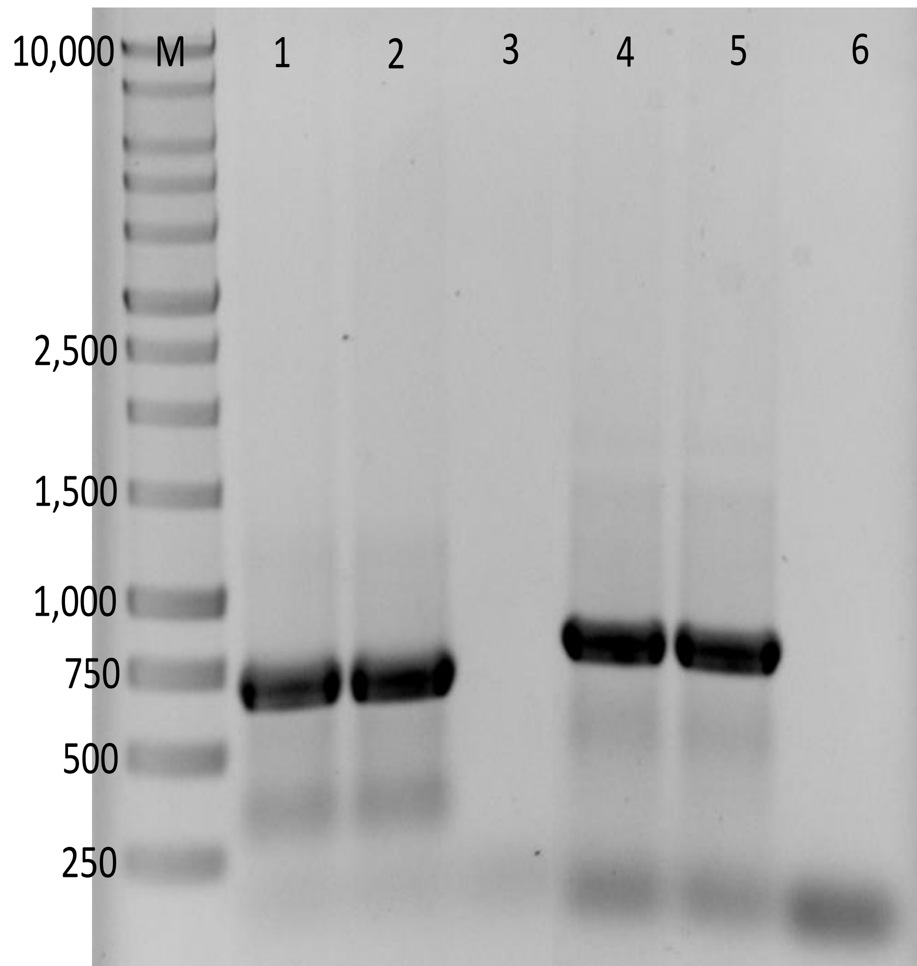


Figure 4.9: Image of agarose gel electrophoresis of PCR products following amplifications of the IE and OE junctions of mTn*5*SS*gusA*31 in the genome of CIAT 899-GE. Lanes 1 and 2 are loaded with the products of amplifying the IE junction with the primer pair IE-F and IE-R. Lanes 3 and 6 are reactions of the negative controls. Lanes 4 and 5 are loaded with the products of amplifying the OE junction with the primer pair OE-F and OE-R. M=1 kb DNA ladder (Promega)

Sequences of the products from the above amplifications gave the expected hits in BLAST, thereby confirming the insertion point of mTn*5*SS*gusA*31 in the genome of CIAT 899-GE. A reconstruction of the insertion point is shown in Figure 4.10.

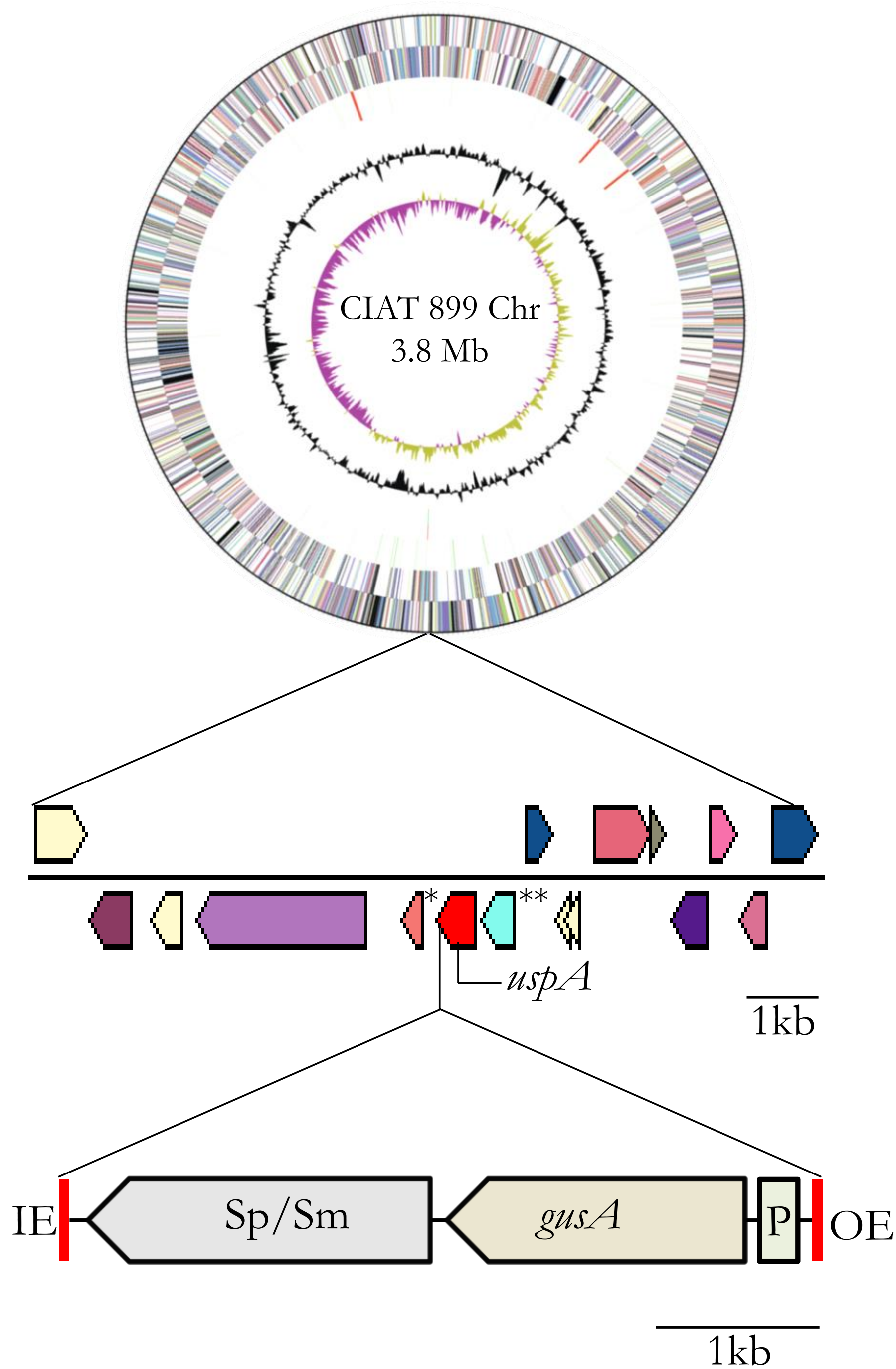


Figure 4.10: A schematic representation of the insertion point of mTn*5*SS*gusA*31 in the chromosome of *R. tropici* CIAT 899-GE. Insertion occurred 6 bp upstream of the universal stress protein A gene (*uspA*) stop codon. *uspA* neighbourhood is shown as well as the map of the transposon. The symbiotically active P*nifH* is labelled P. Map of mTn*5*SS*gusA*31 adapted from that of Wilson et al. (1995). *uspA* locus tag: RTCIAT899_CH07670

*Xanthine phosphoribosyltransferase **molybdenum cofactor synthesis domain-containing protein

4.3.2 Stable broad-host-range CelB plasmid (pGM01) expressed in rhizobia

4.3.2.1 Construction of pGM01

The CelB-mNeonGreen fragment was synthesized and cloned into pMK-RQ to form a 4, 647 bp pMK-RQ_celBmNG. Restriction digestion of pMK-RQ_celBmNG with XbaI and PstI released a 2,337 bp celB-mNeonGreen fragment and other smaller fragments (Figure 4.11). The celB-mNeonGreen fragment was successfully purified from the gel and ligated to the pJP2 backbone, following the excision of *iudA* with XbaI and PstI (Figure 4.11), to form a 12,736 bp pGM01 (Figure 4.12).

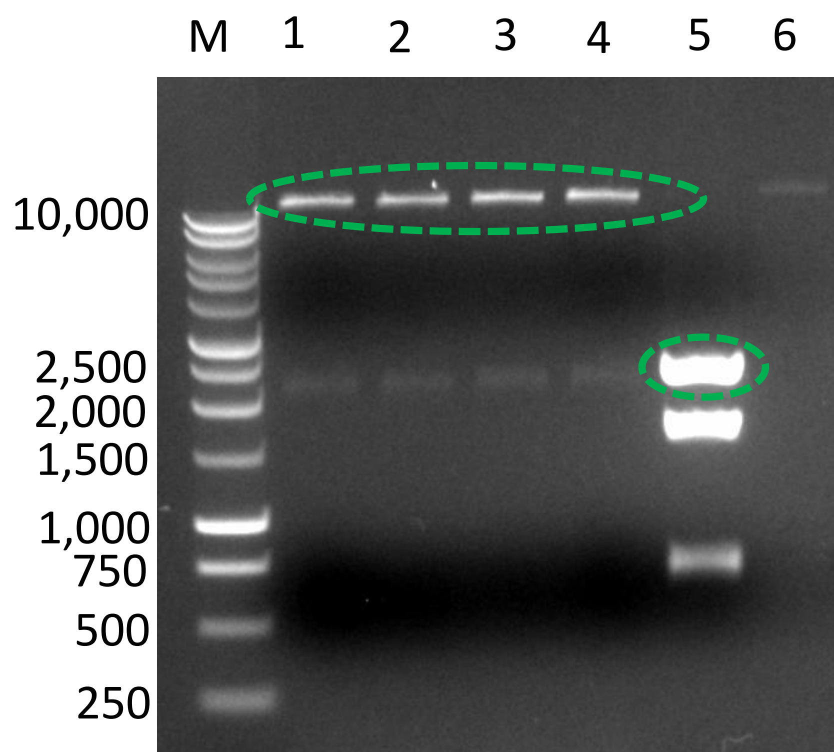


Figure 4.11: Gel image of electrophoresed products of restriction digestion of pJP2 (Lanes 1 to 4) and pMK-RQ_celB-mNG (Lane 5) with XbaI and PstI. Lane 6 was loaded with uncut pJP2. Products circled in green were excised and ligated together to form pGM01. M=1 kb DNA ladder (Promega)

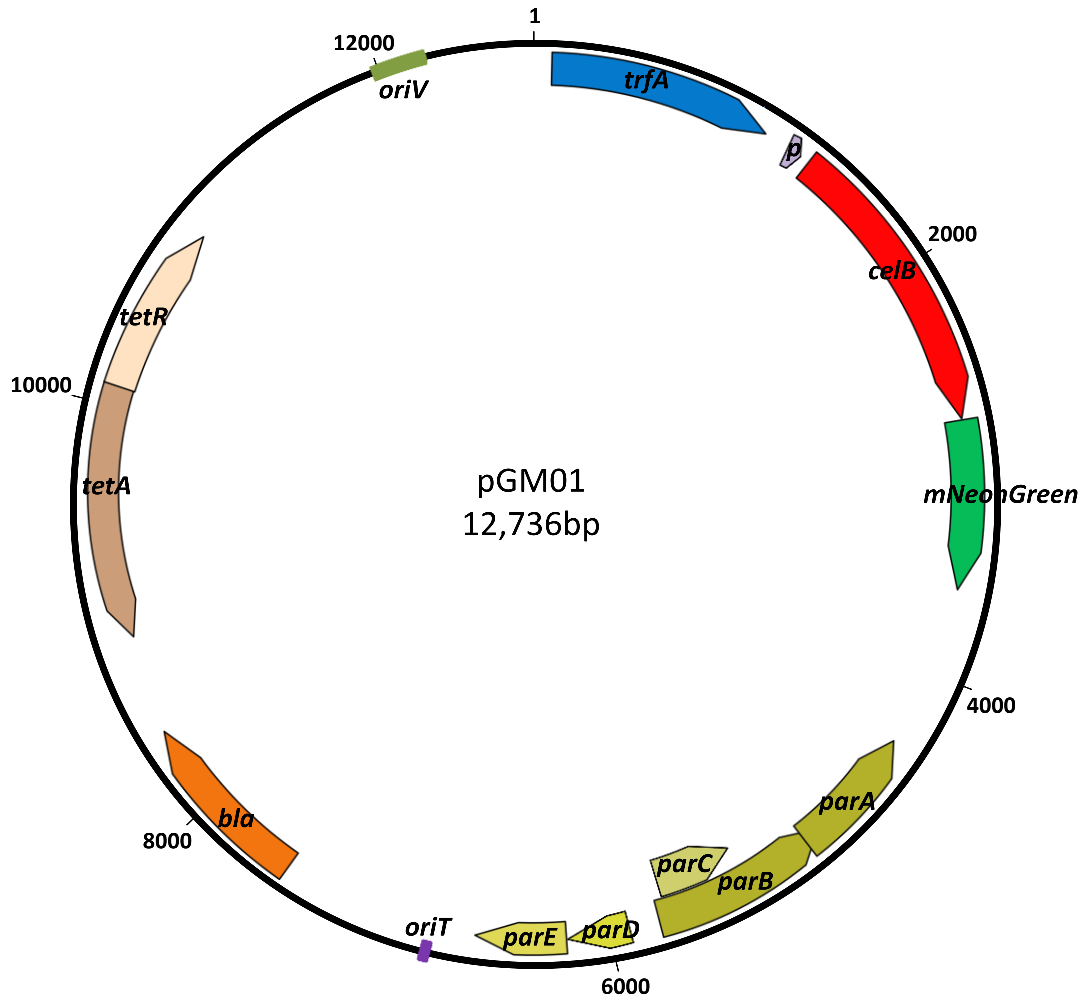


Figure 4.12: Map of pGM01 plasmid (12,736bp). pGM01 has a pJP2 backbone (Prell et al., 2002) carrying the *par* stability region, replication protein gene, ampicillin and tetracycline resistance genes. Additionally, pGM01 carries genes coding for CelB and mNeonGreen protein. Note: The green fluorescent protein, mNeonGreen, was poorly expressed in the vector and was therefore not used as a marker in any work presented in this thesis.

4.3.2.2 Conjugation of pGM01 into rhizobia and expression of thermostable β-glucosidase

The plasmid pGM01 was successfully transformed into *E. coli* ST18 and conjugated into 39 strains of rhizobia by bi-parental mating. Unlike the wild types, transconjugants grew in media containing tetracycline. The presence of the plasmid in the transconjugants was further confirmed by the amplification of a 589 bp region using the primer pair PGM9459F and PGM1537R (primer pair internal of *celB*) (Figure 4.13).

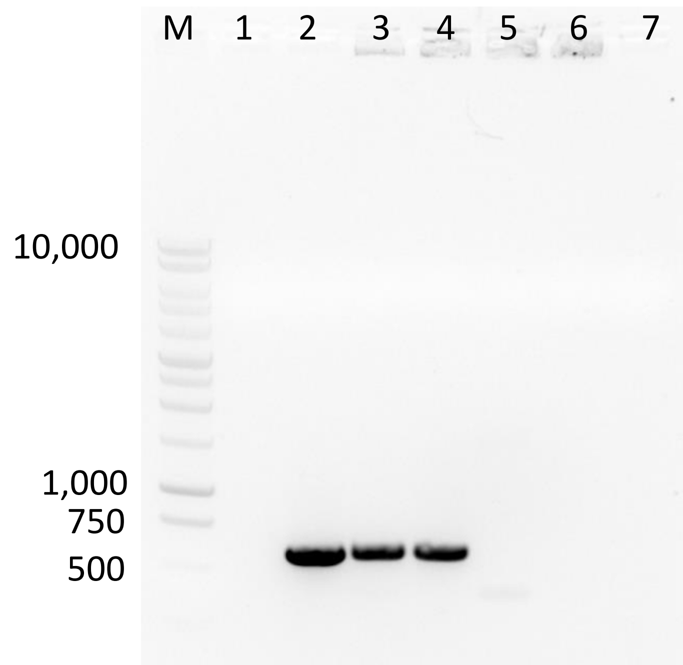


Figure 4.13: Gel electrophoresis image of PCR products with PGM949F and PGM1537R primer pair. Lane 1: pJP2 as a control, Lane 2: pGM01 as a positive control, Lane 3: NAK 103C*, Lane 4: 8002C*, Lane 5: NAK 103, Lane 6: 8002, and Lane 7: No DNA. Templates containing the *celB* region show a strongly amplified 589bp fragment (Lane B, C, and D). Marker: 1 kb Promega ladder.

Additionally, heat treating transconjugants by incubation at 70ºC for 45 min before the addition of magenta-Gal, confirmed the product of the engineered *celB* to be heat stable, as, in contrast to the transconjugants, wild-types did not show any β-galactosidase activity following heat treatment (Figure 4.14).

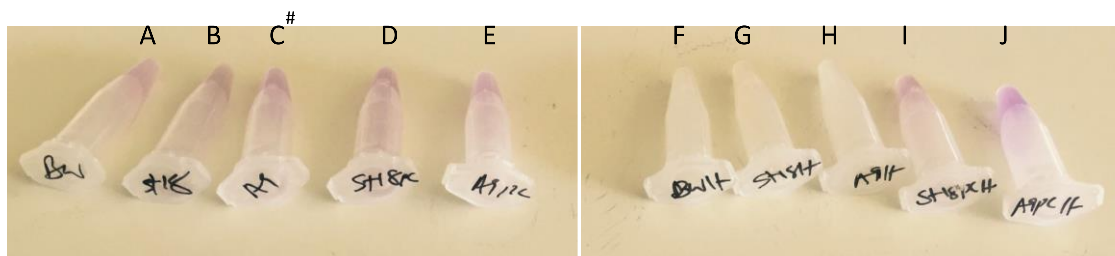


Figure 4.14: β-galactosidase activity (indicated by magenta colour) of (**A**) *E. coli* BW27067 (**B**) *E. coli* ST18 (**C**) NAK 103 (**D**) *E. coli* ST18C and (**E**)NAK 103C cell lysates following the addition of Magenta-gal. Tubes F, G, H, I and J are cell lysates of strains in A, B, C, D and E, in corresponding order, but with Magenta-gal added following heat treatment. The two transconjugants, *E. coli* ST18C and NAK 103C, in tubes I and J respectively, retain β-galactosidase activity after heat treatment.

4.3.2.3 Effect of pGM01 on growth rate and morphology of rhizobia

Rhizobia carrying pGM01 were compared by colony characteristics to their wild types. On both selective and non-selective media, the transconjugants were indistinguishable from their unmarked wild-types by characteristics such as colony size, colour, elevation and mucoidness (Figure 4.15).

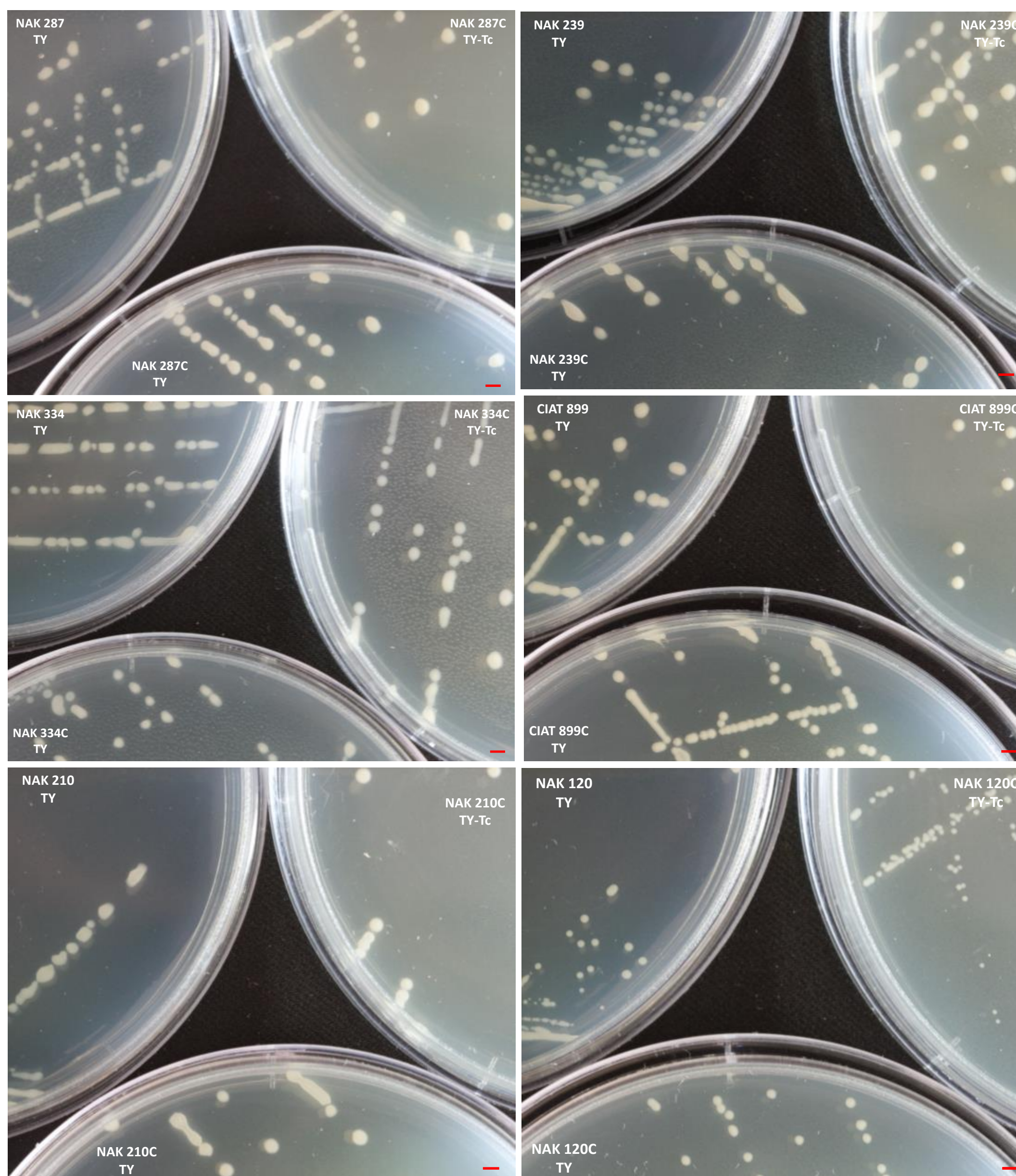


Figure 4.15: Colonies of six strains (NAK 287, NAK 334, NAK 210, NAK 239, CIAT 899 and NAK 120) and their *celB* carrying-pGM01 transconjugants (NAK 287C, NAK 334C, NAK 210C, NAK 239C, CIAT 899C and NAK 120C) after 4 d on TY agar plates with or without tetracycline at 28ºC. Each photo panel contains one wild-type strain on plain TY and it's transconjugant on plain TY and TY with tetracycline. No morphological differences were noted between transconjugants and their unmarked wild types. Scale bar is 2mm.

To further test for similarities or differences between the marked strains and their unmarked wild-types, growth rates were determined, by spectrophotometric means, for a subset of seven strains. The growth curves were highly similar (Figure 4.16) and the mean generation times were similar between the transconjugants and the wild-types ($p>0.05$).

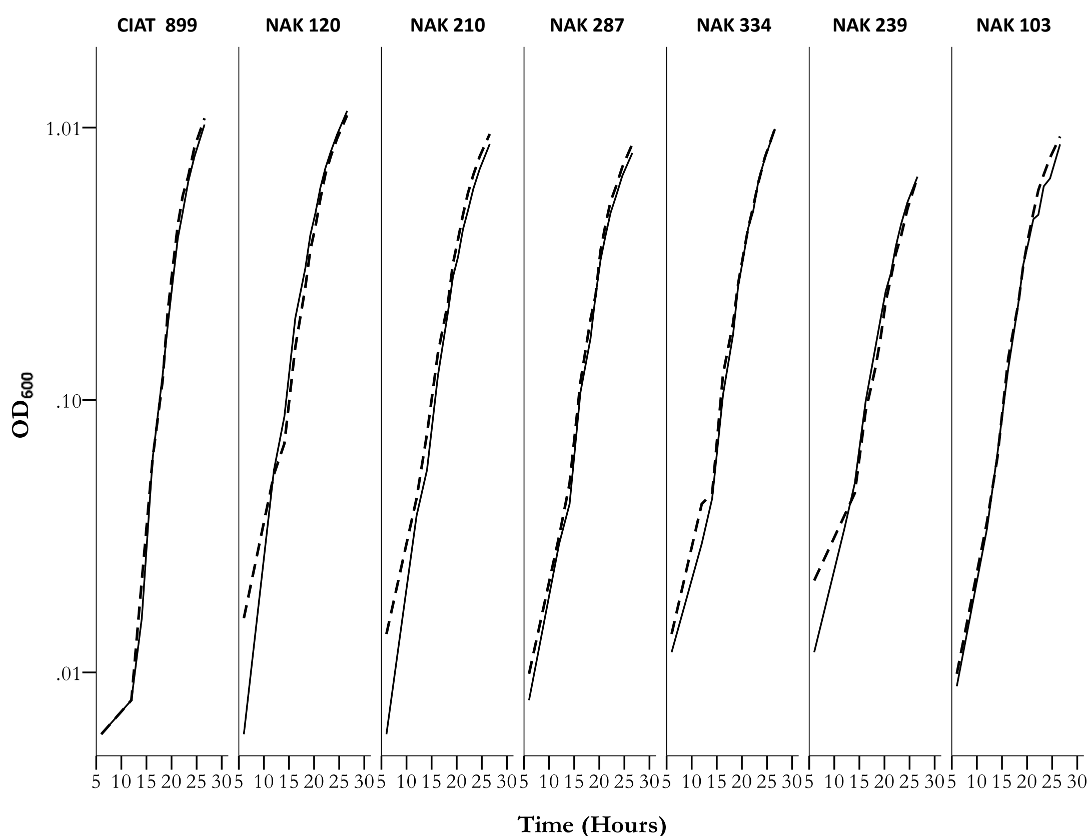


Figure 4.16: Graphs showing truncated growth curves of strains CIAT 899, NAK 120, NAK 210, NAK 287, NAK 334, NAK 239 and NAK 103 carrying pGM1 (---) and their wild-types (—). The wild-types and their transconjugants were grown in non-selective 20 mL TY broth at 28°C with aeration (220 rpm), and OD_{600nm} determined 14 times over a 27 h period.

4.3.2.4 Stability of pGM01 in rhizobia in the absence of antibiotic selection

Three transconjugants carrying pGM01 were assayed for their ability to retain the pGM01 plasmid in the absence of antibiotic selection over approximately 80 generations. No loss of pGM01 was observed in NAK 120C and NAK 210C over the 80 generations (Figure 4.17). In the absence of selection, and in contrast to NAK 120C and NAK 210C, NAK 334C lost pGM01 and after 22 generations, only approximately

50% of cells still retained pGM01. The assay was terminated at approximately 80 generations, at which point only 7% of NAK 334C cells still harboured pGM01 (Figure 4.17).

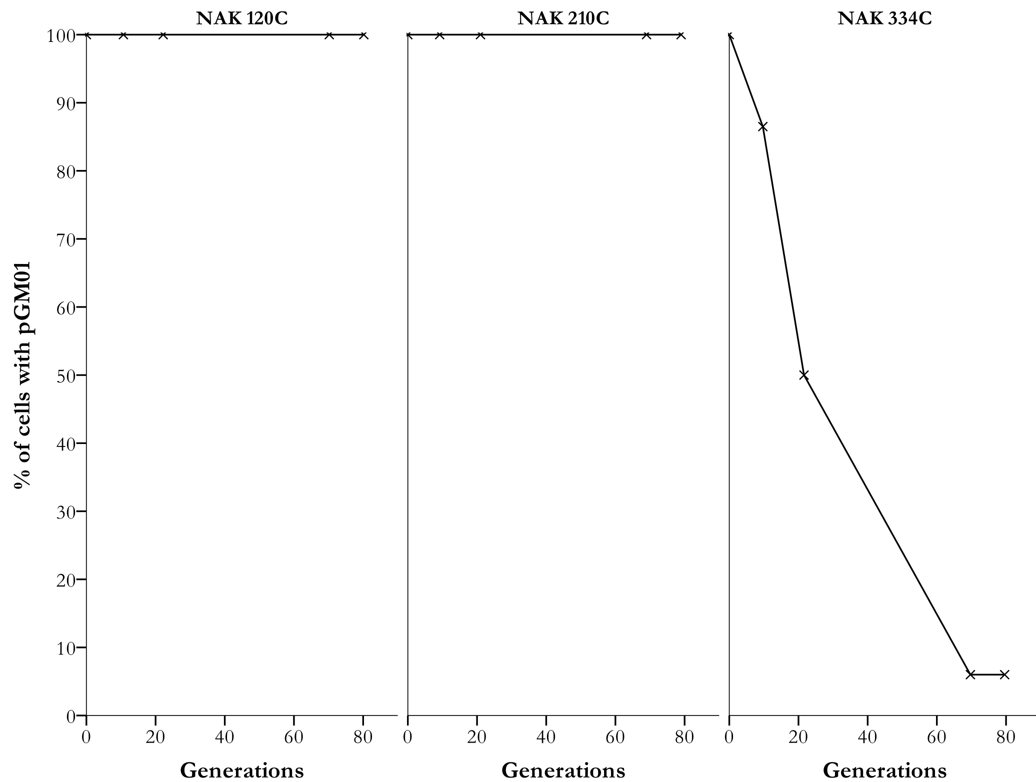


Figure 4.17: Maintenance of pGM01 plasmid in strains NAK 120, NAK 210 and NAK 334 in the absence of antibiotic selection. Strains were grown to mid-log phase in TY broth with tetracycline and sub-cultured in TY broth without antibiotic. The culture was allowed to grow to mid-log and then diluted. This was done for eight consecutive times. Differences in viable counts at start and end time points were used to calculate generations elapsed while the proportion of cells with pGM01 at each time point was determined by replica patching 100 colonies onto TY plates with tetracycline.

4.3.3 Competitiveness of strains against CIAT 899-GE for nodulation of *P. vulgaris*

Nodule occupancy was determined by visual observation of the nodules following staining of roots on all but four treatments that had poor plant growth. Nodules occupied by CIAT 899-GE were blue while those occupied by strain carrying pGM01 were magenta (Figure 4.18). Occasionally, nodules stained both blue and magenta,

indicating dual infections (Figure 4.19). Additionally, some nodules were either fully unstained or had unstained portions (Figure 4.20). Inoculation with seven strains resulted in unstained or partially stained nodule phenotypes, representing 18% of strains tested. The seven were NAK 103, NAK 287, NAK 299, NAK 315, NAK 334, NAK 368 and NAK 407. In the calculation of competitiveness indices, fully unstained nodules were scored as magenta as these are likely to have arisen from nodulation by the *celB*-marked strains as no nodules were found in the un-inoculated treatments and, additionally, no unstained nodules were seen when plants were only inoculated with CIAT 899-GE. Re-isolation from the unstained nodules was not possible because roots had been heat-treated as part of the staining process. Dual occupied nodules, as indicated by portions of magenta and blue or unstained and blue, were omitted from consideration in the computation of the competitiveness of the strains.

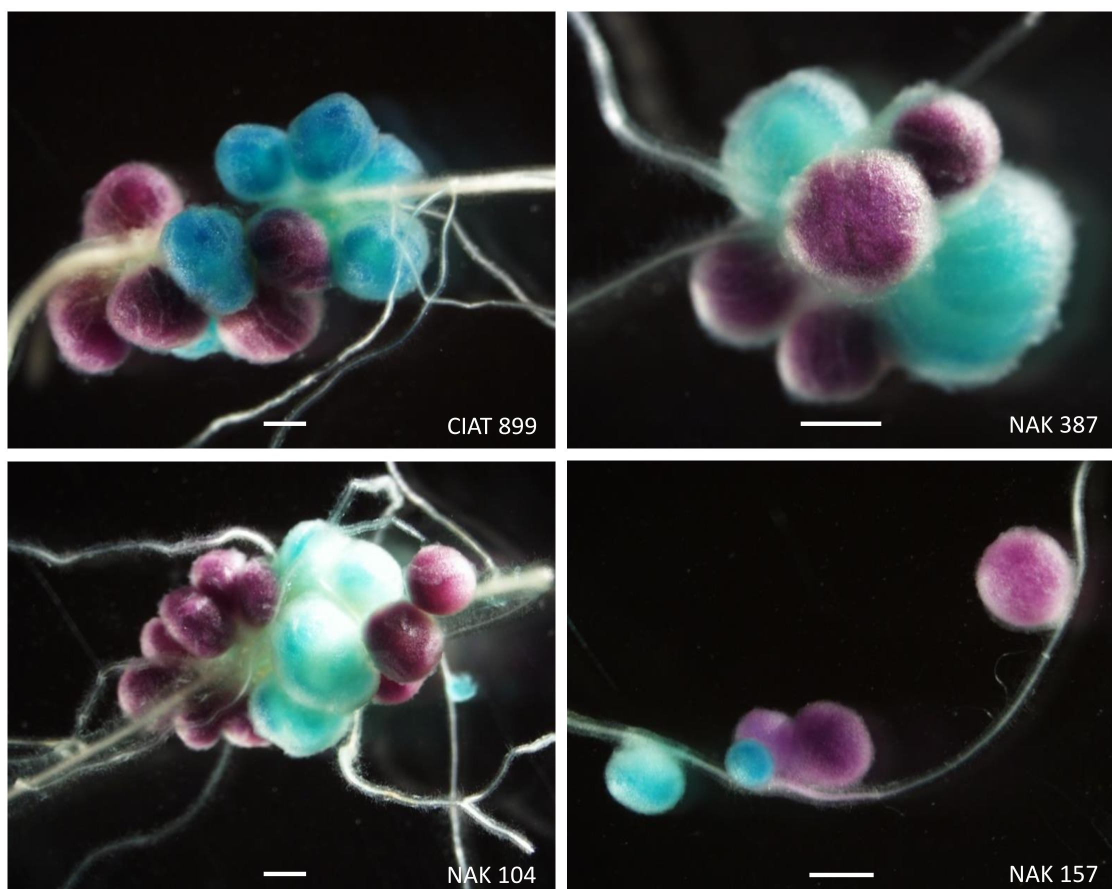


Figure 4.18: *P. vulgaris* root nodules following co-inoculation with a *gusA* marked *R. tropici* CIAT 899-GE (blue nodules) and strains carrying the *celB* encoding plasmid, pGM01 (magenta nodules). The identity of the pGM01-carrying strain is indicated on each photo. Scale bar is 1 cm.

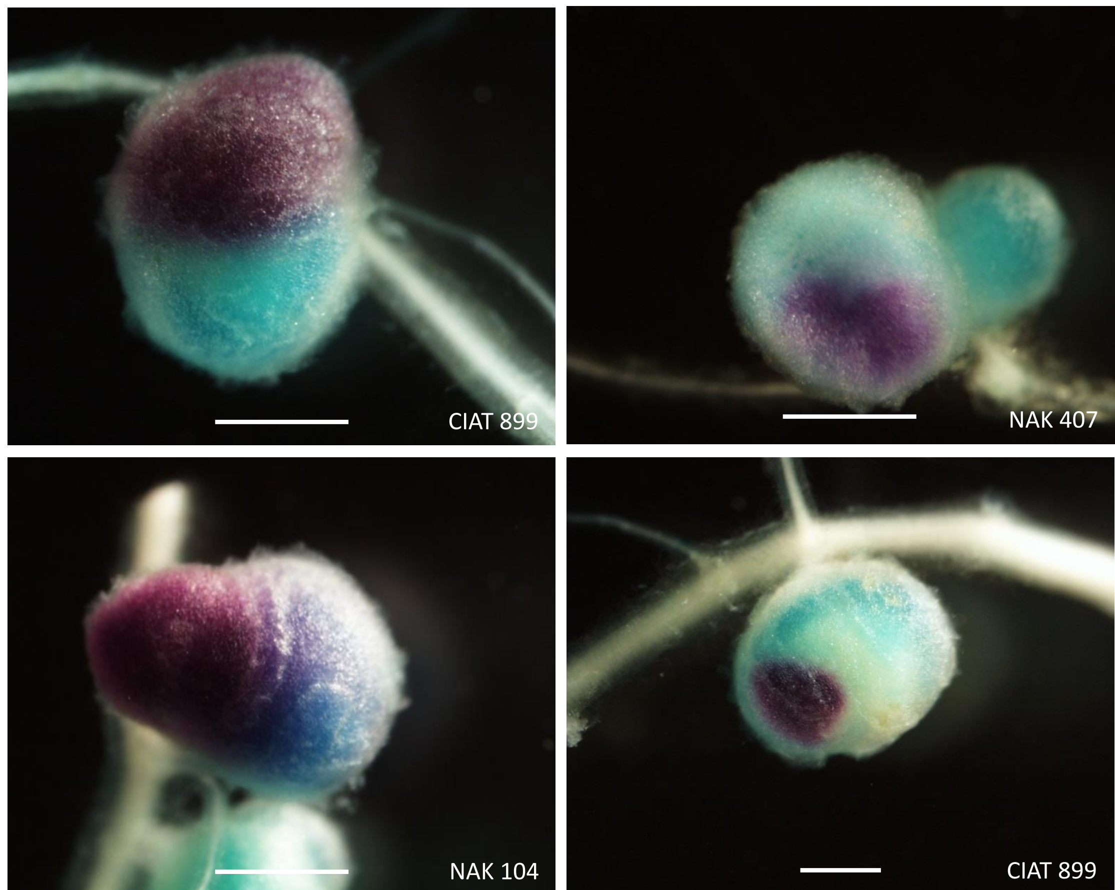


Figure 4.19: Dual nodule occupancy in *P. vulgaris* root nodules following co-inoculation with a *gusA* marked *R. tropici* CIAT 899-GE (blue portions) and strains carrying the *celB* encoding plasmid, pGM01 (magenta portions). The identity of the pGM01-carrying strain is indicated on each photo. Scale bar is 1 cm.

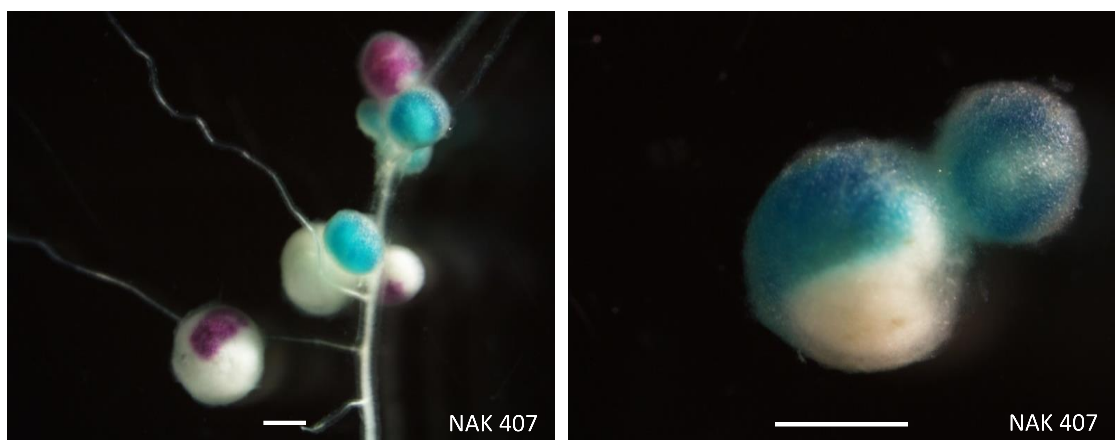


Figure 4.20: Fully unstained or partially unstained *P. vulgaris* root nodules following co-inoculation with a *gusA* marked *R. tropici* CIAT 899-GE and strain NAK 407, carrying the *celB* encoding plasmid, pGM01. Scale bar is 1 cm.

The competitiveness of co-inoculated strains was calculated as described in Section 4.2.2.3 (Eq. 4.2). CIAT 899 marked with *celB* (CIAT 899C) was found to be of equal competitiveness to CIAT 899 marked with *gusA* (CIAT 899-GE) (Figure 4.21). A competitiveness value of 1 is indicative of competitiveness equal to that of CIAT 899-GE. Seven (7) strains were found to out-compete CIAT 899-GE for nodulation of *P. vulgaris* while 17 strains were out-competed by CIAT 899-GE (LSD, $P<0.05$) (Figure 4.21). Some of the more competitive strains were NAK 103, NAK 287, NAK 315 and NAK 277. Very uncompetitive strains included *Rhizobium* sp. strain 8002, NAK 303 and NAK 367.

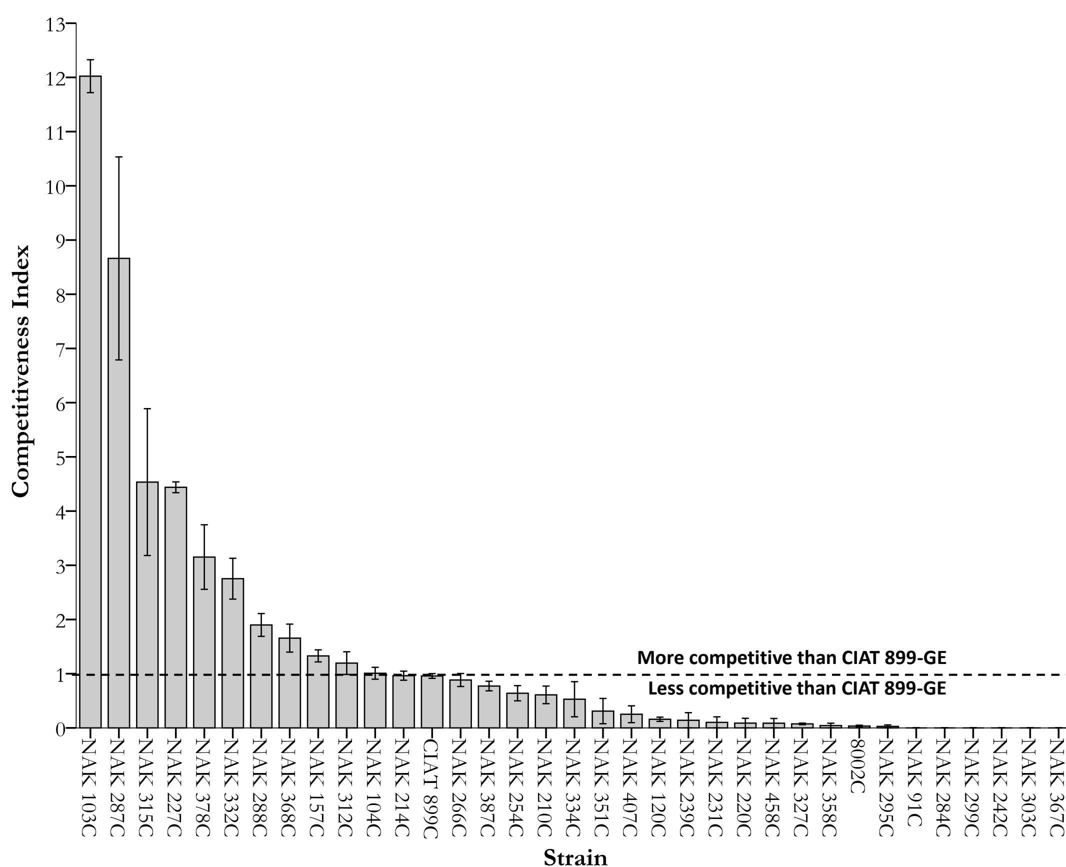


Figure 4.21: Graph showing the competitiveness of 35 strains for nodule occupancy of *P vulgaris* cv. KK08 nodules, against CIAT 899-GE. Columns are means of 2-6 plant replicates. Bar ± 1 standard error of mean. Competitiveness Index (CI) is defined as test strain (carrying pGM01)-to-CIAT 899-GE ratio in the nodules divided by test strain-to-CIAT 899-GE ratio in the inoculum. A CI of 1 indicates equal competitiveness to CIAT 899-GE.

4.4 Discussion

4.4.1 Dual marker system

The work presented in this chapter succeeded in developing and evaluating pGM01, an IncPα broad-host-range plasmid carrying *celB* marker gene, for suitability in marking diverse rhizobia. pGM01 was found to be stably maintained in most the rhizobial strains tested, both *in vitro* and in nodules. The plasmid was rapidly mobilizable into rhizobia by bi-parental conjugation, facilitating the screening of a large number of strains for competitiveness in nodulating *P. vulgaris*.

In this study, the *celB* gene on the pGM01 plasmid was expressed in free-living and symbiotic rhizobia and importantly, all tested marked strains were not altered in their growth characteristics or their competitiveness for nodulation of *P. vulgaris* (Figures 4.15, 4.16 and 4.21). The lack of a detrimental phenotype was in agreement with observations by Duodu et al. (2008) who noted no detrimental effects on growth, survival and nodulation of tagging *R. leguminosarum* bv. trifolii strains with plasmid-borne GFP or DsRed. Although Corich et al. (2001) reported the formation of tiny nodules by strains carrying a plasmid-borne *lacZ* marker, the phenotype is likely attributable to an over-expression of the *lacZ* gene under the strong synthetic promoter (P*syn*). In pGM01, the *celB* expression is driven by P*tac*, a constitutive promoter that is weakly expressed in *Rhizobium* spp. (Giacomini et al., 1994) and this may explain the lack of any atypical nodule phenotypes or aberrant growth rates indicative of a plasmid-related energy drain on the marked strains.

The broad host range nature of pGM01 was confirmed by the successful transfer, expression and maintenance of the plasmid in diverse genetic backgrounds that included *R. tropici* (CIAT 899), *R. sophoriradicis* (NAK 387), *R. phaseoli* (NAK 458, NAK 287) *R. paranaense* (NAK 120), *R. etli* (NAK 157) and others of potentially undescribed taxa (NAK 266, NAK 210, NAK 312, NAK 358) (Figure 4.21). However, despite pGM01 having an RK2 backbone belonging to the IncPα family of plasmids shown to be stably maintained in a wide range of Gram-negative bacteria (Easter et al., 1998), plasmid loss was observed in some strains. For example, NAK 334C lost pGM01 when grown *in vitro* in the absence of antibiotic selection (Figure 4.17). Unstained nodules (Figure 4.20) were observed following inoculation of plants with seven of the 35 marked strains, including NAK 334C, indicating a loss of plasmids in these strains.

Loss of introduced plasmids has been observed before in rhizobia (Corich et al., 2001; Duodu et al., 2008; Gage et al., 1996) but instability in rhizobia of plasmids bearing the RK2 backbone is unknown. The RK2 backbone encodes two stability mechanisms that had previously demonstrated to be efficient in a narrow range of rhizobia (Prell et al., 2002; Weinstein et al., 1992). The stability mechanisms are encoded by the *parCBA/DE*. The *parCBA* operon encodes an active partitioning system that distributes plasmids within the cell to ensure daughter cells receive a plasmid upon division (Easter et al., 1998; Sobecky et al., 1996) while *parDE* is responsible for post-segregational killing relating to a toxin and an antitoxin (TA) system. The reason for pGM01 instability in some rhizobial backgrounds will need to be investigated. One possibility is a degradation of the toxin-antitoxin (TA) stability system encoded by pGM01 through cross-reaction with a homologous or different TA system naturally encoded by some strains (Goeders & Van Melderen, 2014). In spite of the instability observed with some strains, pGM01 was stable in >80% of the strains tested.

For the simultaneous detection of nodule occupants in *P. vulgaris* during competition studies, strains carrying pGM01 were co-inoculated with CIAT 899 marked with *gusA* using a mini-transposon. The aim of the rhizobial competition studies was to investigate the extent to which the individual Kenyan strains limit the occupancy of *P. vulgaris* nodules by CIAT 899 in Kenyan soils. Therefore, the chromosomal marking of CIAT 899 using a mini-transposon was found suitable as only one reference strain needed to be marked.

From this work, two CIAT 899 mutants with a chromosomal *gusA* marker gene are now available. In CIAT 899-GE, the *gusA* is expressed from a symbiotically active P*nifH* that, as previously reported (Wilson et al., 1995), resulted in strong *gusA* expression localized in nodules (Figure 4.6C). The mTn*5*SS*gusA*31 in CIAT 899-GE was found inserted in the universal stress protein A gene (*uspA*) (Figure 4.10). The exact biochemical role of *uspA* in bacteria is unknown but it is upregulated in stress conditions such as exposure to UV light, carbon starvation, heat shock and metals (Kvint et al., 2003). Two other genes in the annotated CIAT 899 genome (GenBank: PC40615) are designated putative *uspA* genes (loci tags RTCIAT899_CH02385 and RTCIAT899_PC04615), but a low amino acid identity of the two to the *uspA* disrupted in CIAT 899-GE suggested the three genes are not functionally equivalent.

When subcultured, the *uspA* insertion mutant had a considerably longer lag phase than was seen with the wild-type but after the lag phase, the mutant grew at a rate indistinguishable from that of the wild-type (Figure 4.5). A similar growth phenotype was observed in *E. coli* defective in UspA synthesis and was linked to a possible role of UspA in the flow of carbon in the central metabolic pathways of *E. coli* (Nyström & Neidhardt, 1993). However, CIAT 899-GE (*uspA*::mTn*5*-Sp/Sm-*gusA*) was unaltered in nodulation and competitiveness, thus the gene is unlikely to be involved in symbiotic N_2 fixation in *P. vulgaris*.

4.4.2 Competitiveness of strains against CIAT 899

CIAT 899 was found to be more competitive than a majority of the 34 strains isolated from Kenyan soils, with only approximately 20% of the strains outcompeting CIAT 899.

Subsequent analysis involving effectiveness data (Chapter 3) and the competitiveness data revealed a weak positive correlation between effectiveness and competitiveness of strains (R^2=0.23). To illustrate this, NAK 312 (poorly effective) and NAK 287 (effective) were at least as competitive as CIAT 899 while NAK 210 (ineffective) and NAK 458 (effective) were less competitive than CIAT 899. Strains that were both competitive and effective and, therefore, offer promise as future *P. vulgaris* inoculant strains include NAK 227, NAK 288, NAK 287, NAK 157, NAK 104, and NAK 214.

CIAT 899 is the major *P. vulgaris* inoculant in Kenya (Bala et al., 2011). Therefore, findings in this chapter, in addition to revealing potential inoculant strains, reveal the rhizobial genotypes likely to be permissive or restrictive to the successful nodulation of *P. vulgaris* by CIAT 899 in Kenyan soils.

4.4.3 Concluding remarks

Work in the chapter developed and tested pGM01, a broad-host-range plasmid carrying *celB* that was found suitable for use in rhizobial competition studies. The introduction of the plasmid into rhizobia was rapid. For the majority of the strains, the plasmid was expressed and maintained stably *in vitro* and in nodules. Consequently, pGM01 is a useful tool for nodule occupancy studies. pGM01 was combined with a second marker borne on a mini-transposon to successfully screen 34 rhizobial strains isolated from Kenyan soils for occupancy of *P. vulgaris* nodules in a competitive scenario.

However, the competition experiments conducted involved the application of competing strains as liquid cultures on germinating seeds, and while informative as a preliminary analysis, do not mirror competition scenarios as they occur in the field following inoculation. Combining pre-inoculated soil and seed inoculation may be a better way to mimic competition under field settings and therefore arrive at more accurate predictions of the competitive nature of strains.

CHAPTER 5
The determinants of inoculation success in *P. vulgaris*

5.1 Introduction

The inoculation of *P. vulgaris* with rhizobia, to boost N_2 fixation and enhance yields, is a common practice in many parts of the world. However, the outcomes of inoculation vary greatly, with success and failure reported (Giller et al., 1998; Hungria et al., 2000; van der Bom, 2012). This variability in response to inoculation is largely due to the presence of indigenous rhizobia that compete with the inoculant for nodule initiation sites on roots of *P. vulgaris* (Thies et al., 1991b).

A considerable amount of research effort has been expended to understand rhizobial competition, especially towards elucidating the role of population densities of indigenous rhizobia in determining responses to inoculation. The population density of soil rhizobia is believed to be inversely related to the inoculation response (Thies et al., 1991b). However, there is no unanimity in literature on the density or nature of indigenous rhizobia capable of inhibiting a response to inoculation. For example, while as few as 93 or 700 rhizobia per g of soil can present an insurmountable barrier to successful inoculation of *P. vulgaris* (Thies et al., 1991a; Vargas et al., 2000), inoculation response has been achieved in soils with 10^3 to 10^5 rhizobia per g of soil (Hungria et al., 2000; Hungria et al., 2003; Vlassak et al., 1996). These variations in responses at different rhizobial densities indicate that other factors, beside the population density of indigenous rhizobia, strongly influence the outcomes of inoculating *P. vulgaris*.

One possible factor is the genotype(s) of the background rhizobia. Rhizobia differ in many attributes such growth rates, carbon-source preferences, ability to colonise the rhizosphere of legumes and affinities towards legume hosts (Aguilar et al., 2004; Yates et al., 2008). Such strain differences are expected to lead to variable inoculation outcomes. However, no known studies have investigated the importance of different background rhizobial genotypes to the outcomes of inoculating *P. vulgaris*.

Inoculant quality is a second factor that might result in variable responses to the inoculation of *P. vulgaris* (Graham & Ranalli, 1997). A good quality inoculant contains at least 10^9 cells of rhizobia g^{-1} carrier and delivers at least 10^5 rhizobia per seed to large seeded legumes (Catroux et al., 2001a; Hungria et al., 2005; Lupwayi et al., 2000). In some instances, the population threshold is not maintained throughout the shelf-life of the inoculants (Balume et al., 2015) leading to suboptimal numbers being applied onto seeds. As was shown by Hume and Blair (1992) on soybean, applying a suboptimal

inoculum dosage onto a specifically nodulating legume in soils with few or no compatible rhizobia leads to a greatly reduced response to inoculation. *P. vulgaris* nodulates promiscuously and is commonly grown in soils with 10^4-10^6 rhizobia per g of soil (Alberton et al., 2006; Andrade et al., 1999; Anyango et al., 1995; Hungria et al., 2000; Kawaka et al., 2014; Langwerden, 2014) and the role of inocula dosage in determining response to inoculation in high rhizobial backgrounds is unclear.

A third factor that might influence inoculation outcomes is soil mineral N. In addition to the potential of soil N to suppress or inhibit nodulation and N_2 fixation (Hartwig, 1998), soil N can alter the nodule occupancy outcomes in multi-strain environments. The alteration in nodule occupancies is believed to be either due to differences in nitrate tolerance by the strains (Vargas et al., 2000) or the production of altered root exudates by the host that interact with the symbioses (Caballero-Mellado & Martinez-Romero, 1999). The effect of N is, however, equivocal as other studies have found soil N to have no role in determining the infection success of one strain over another (Abaidoo et al., 1990; Glyan'ko et al., 2009). Farmers grow and inoculate *P. vulgaris* in soils that vary in N and knowledge on the influence of soil N on the establishment and success of inoculants would be valuable.

Studies in this chapter aimed to improve the understanding of the outcomes of *P. vulgaris* inoculation by investigating the role of the rhizobial genotype, background rhizobial population size, inoculant dosage and mineral N in the nodule occupancy outcomes of CIAT 899, in the presence of high soil numbers of *R. paranaense* NAK 120, *Rhizobium* sp. NAK 210 and *R. phaseoli* NAK 287, all isolated from Kenyan soils. The study applied the dual marker system developed in Chapter 4 to identify nodule occupants and to enumerate rhizobia at different stages of the two experiments conducted.

5.2 Materials and Methods

5.2.1 Bacterial strains and *P. vulgaris* cultivar

Studies in this chapter used CIAT 899-GE, CIAT 899C, and three strains from Kenyan soils viz., *R. paranaense* NAK 120C, *Rhizobium* sp. NAK 210C and *R. phaseoli* NAK 287C. CIAT 899-GE carries genes for *gusA* and resistance to spectinomycin (Section 4.3.1). NAK 120C, NAK 210C, NAK 287C and CIAT 899C carry pGM01 which codes for *celB* and tetracycline resistance (Section 4.3.2). Marked strains were as competitive and fixed similar amounts of N as parental strains (Section 4.3.1.2; 4.33). In Experiment 2 of this chapter, CIAT 899-GE was used as the seed inoculant and the rest were, independently, used as soil inocula. The 'soil strains' differed in their relative effectiveness and competitiveness (Figure 5.1) and were therefore suitable in evaluating the effect of different soil rhizobial genotypes on the nodule occupancy of a seed-applied inoculant strain (CIAT 899-GE). All experiments were conducted with *P. vulgaris* cv. Kenya Tamu.

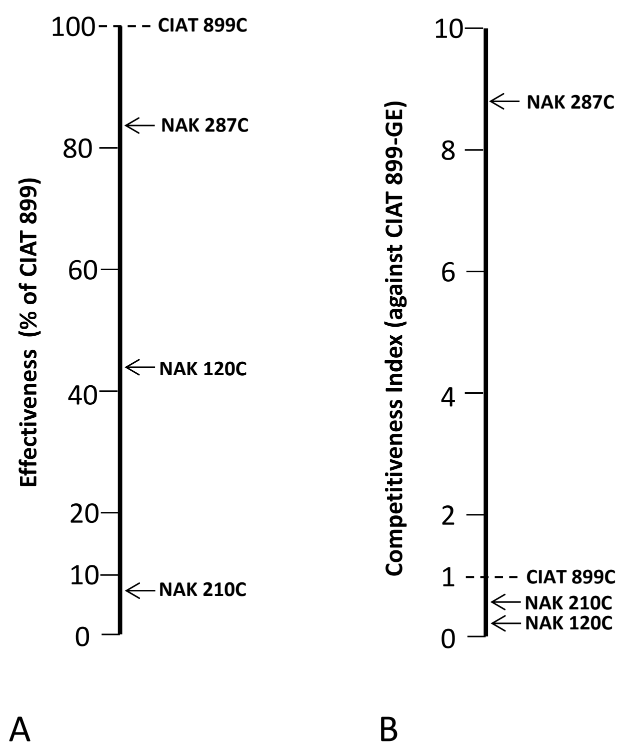


Figure 5.1: The relative effectiveness (**A**) and competitiveness (**B**) of NAK 120C, NAK 210C, NAK 287C and CIAT 899C. Effectiveness was calculated by expressing shoot dry weights (SDW) as a percentage of SDW of CIAT 899. Competitiveness Index was determined from liquid coinoculations by dividing strain-to-CIAT 899-GE ratio in the nodules by strain-to-CIAT 899-GE ratio in the inoculum.

5.2.2 Experiment 1: Rhizosphere competence of strains

Bacterial strains differ in their ability to survive and multiply in the rhizosphere of plants and a preliminary experiment was carried out to determine whether differences in rhizosphere competence were responsible for differences in competitiveness (for nodulation) of the four rhizobial genotypes. To do this, NAK 287C, NAK 210C and NAK 120C were co-inoculated with CIAT 899-GE onto *P. vulgaris* at a low density as liquid inoculum and the proportion of each strain in the rhizosphere and in nodules subsequently determined.

Inoculations

The five strains were grown in TY broth with appropriate antibiotic selection and suspensions of 2×10^3 cells mL^{-1} made in sterile deionised water (Section 4.2.4). Suspensions were mixed in a 1:1 ratio to give four combinations viz.; CIAT 899C/CIAT 899-GE, NAK 120C/CIAT 899-GE, NAK 210C/CIAT 899-GE and NAK 287C/CIAT 899-GE. For each combination, a 1 mL aliquot was removed and used for viable cell counts by the Miles and Misra method (O'Hara et al., 2016) on TY (plain TY, TY with spectinomycin (200 µg mL^{-1}), and TY with tetracycline (20 µg mL^{-1}) and 1 mL of the remainder applied per pre-germinated seed at sowing in 3.5 L pots filled with a steam-sterilized sandy soil (2 ppm N, 0.08% (w/w) organic C, pH ($CaCl_2$) 7.1). Seeds in the un-inoculated control treatment received 1 mL of sterile de-ionized water per seed. Treatments were replicated in eight pots, each sown with two seeds and randomized in a naturally-lit glasshouse maintained at 22°C. On emergence, plants were thinned to one per pot. All other procedures, including maintenance of plants, were as per Section 3.2.2.

Assessing rhizosphere colonization

Seven days after inoculation, four pots were selected from each treatment, plants carefully removed from the pots and roots gently shaken to remove excess soil. The roots, together with tightly adhering soil, were transferred into pre-weighed 50 mL sterile falcon tubes, weighed, and diluted 10-fold (w:v) with sterile water before shaking for 30 min at medium speed on an Analite wrist shaker (Analite Pty, Australia). Viable cell counts of the resulting suspensions were then determined by the Miles and Misra technique on TY agar, TY with spectinomycin (200 µg mL^{-1}), and TY with tetracycline (20 µg mL^{-1}).

Assessing nodule occupancies

Plants in the remaining pots were harvested 17 d after inoculation and roots double stained with X-Glc and Magenta-Gal (Section 4.2.4). Nodules of each colour were then counted and the competitiveness of each strain, against CIAT 899-GE, for nodulation of the host was calculated (Section 4.2.2.3).

5.2.3 Experiment 2: Evaluating the effect of four factors on nodule occupancy by seed-applied CIAT 899-GE

Experimental design

In total, four factors were evaluated for their effect on nodule occupancy by seed-applied CIAT 899-GE: soil rhizobial genotypes (4), soil population densities (2), seed inocula densities (2), and soil N (2). The effect of supplementary N on nodule occupancy was only tested in soils inoculated with NAK 120C and NAK 210C and an incomplete factorial experiment was designed (Table 5.1). All combinations had four pot replicates, each with two plants. Plants were arranged in randomized blocks in a naturally-lit glasshouse maintained at 22°C.

Inoculation of potted soil to 10^2 and 10^5 rhizobia g^{-1} of dry soil

To facilitate the evaluation of the effect of soil population densities on the nodule occupancy of seed-applied CIAT 899-GE, two distinct populations of NAK 120C, NAK 210C, NAK 287C and CIAT 899C were established. The four strains were grown to mid-log phase in TY broth with appropriate antibiotics and cell suspensions with an OD_{600nm} of 0.2 (approximately 2×10^8 cells mL^{-1}) separately prepared in sterile deionised water following washing of cells (Section 4.2.4). These suspensions were further serially diluted in sterile de-ionized water and approximately 2.72×10^5 and 2.72×10^8 cells separately suspended in 80 mL aliquots of sterile water. These numbers resulted in 10^2 and 10^5 rhizobia per gram of dry soil respectively when inoculated into 3.2 kg sandy soil (2 ppm N, 0.08% (w/w) organic C, pH ($CaCl_2$) 7.1) containing 15% moisture.

The 80 mL bacterial suspensions were aseptically (in laminar flow hood) mixed with the steam-sterilized pot soil in a zip lock bag and the soil returned into the pots. At this point, 2 g of the inoculated soil was removed from each pot and together with a soil sample from an uninoculated pot, used to enumerate bacteria by the Miles & Misra technique on TY medium supplemented with tetracycline (within 2-4 h).

Table 5.1: Combinations in the incomplete factorial design used (Experiment 2)

	FACTORS			
	Soil rhizobial genotype	**Population density (g^{-1} soil)***	**Inoculum rate (CIAT 899-GE seed^{-1})****	**Nitrogen*****
1	**NAK 120C**	10^2	10^4	0
2	NAK 120C	10^2	10^6	0
3	NAK 120C	10^2	10^4	1
4	NAK 120C	10^2	10^6	1
5	NAK 120C	10^5	10^4	0
6	NAK 120C	10^5	10^6	0
7	NAK 120C	10^5	10^4	1
8	NAK 120C	10^5	10^6	1
9	**NAK 210C**	10^2	10^4	0
10	NAK 210C	10^2	10^6	0
11	NAK 210C	10^2	10^4	1
12	NAK 210C	10^2	10^6	1
13	NAK 210C	10^5	10^4	0
14	NAK 210C	10^5	10^6	0
15	NAK 210C	10^5	10^4	1
16	NAK 210C	10^5	10^6	1
17	**NAK 287C**	10^2	10^4	0
18	NAK 287C	10^2	10^6	0
19	NAK 287C	10^5	10^4	0
20	NAK 287C	10^5	10^6	0
21	**CIAT 899C**	10^2	10^4	0
22	CIAT 899C	10^2	10^6	0
23	CIAT 899C	10^5	10^4	0
24	CIAT 899C	10^5	10^6	0
Controls				
25	0	0	0	0
26	0	0	0	2*

*Rhizobial densities at soil inoculation **from use of inoculant carrying 10^7 or 10^9 rhizobia g^{-1} of peat ***0=no nitrogen, 1=1.2 mL of 1 M ammonium nitrate pot^{-1} week^{-1}, 2=2.4 mL of 1 M ammonium nitrate pot^{-1} week^{-1}

Pots containing inoculated soils were then covered with cling wrap and kept in a shaded glasshouse maintained at 22°C for 4 d before sowing.

Preparation of inoculants and seed inoculation

To evaluate the effect of seed inoculation rate on nodule occupancy, two peat inoculants of CIAT 899-GE carrying 10^9 and 10^7 rhizobia g^{-1} peat were prepared and used to inoculate *P. vulgaris* seed.

To prepare the inoculant carrying 10^9 rhizobia g^{-1} peat, 50 mL of a stationary phase culture of CIAT 899-GE was injected into 150 g of sterile peat and cured for 10 d at 28ºC. At the end of the curing period, five 1 g samples were removed from the prepared inoculant and the concentration of rhizobia determined by the Miles & Misra technique. The inoculant carrying 10^7 rhizobia g^{-1} peat was prepared 2 hours before it was required for seed inoculation. 1 g of the cured inoculant carrying 10^9 rhizobia g^{-1} peat was mixed into 99 g of sterile peat removed from a 150 g packet of peat injected with 50 mL sterile deionised water. The population of rhizobia in the new inoculant was assessed as described earlier for the inoculant carrying 10^9 rhizobia g^{-1} peat.

Seeds were inoculated with the two inoculants prepared above to obtain two seed inocula dosages using standard techniques (Yates et al., 2016b). Briefly, 2 parts of 40% (w/v) gum arabic were separately mixed with 1 part of peat carrying 10^9 rhizobia g^{-1} or 10^7 rhizobia g^{-1} peat. To inoculate approximately 200 seeds weighing 132 g with any of the inoculants, the seeds were surface sterilised (Section 2.2.2) and left to dry in the laminar flow hood prior to mixing with 3 mL of the inoculant-sticker slurry in a zip lock bag. Inoculated seed was spread out on sterile paper in a laminar flow hood and left to dry for 2 h. Five seeds from each treatment were randomly selected, set aside and later used for the determination of rhizobia numbers per seed by the Miles and Misra method following a shaking of each seed in 10 mL sterile water for 10 mins on an Analite wrist shaker (Analite Pty, Australia). Of the remaining seeds, two were sown per pot containing soil rhizobia. For the uninoculated treatment and the N-control treatments, surface sterilised seed was sown into steam sterilised soil. Except for the nitrogen treatments described below, plants were maintained with nutrient solutions and sterile deionised water (Section 2.2.2). Three weeks after sowing, the pots were flushed with sterile deionised water to prevent salt accumulation in pots and application of nutrients resumed. Plants were harvested 30 d after sowing, roots separated from the shoots,

rinsed then double stained with X-Glc and Magenta-Gal (Section 4.2.4) before counting of nodules by colour.

Nitrogen treatments

To assess the effect of soil N on nodule occupancies, two levels of N were maintained in the NAK 120C and NAK 210C treatments. Pots either received no N or received 0.6 mL of 1 M ammonium nitrate per pot twice per week. The uninoculated N-controls received 1.2 mL of ammonium nitrate per pot twice per week.

Monitoring of rhizobia populations in potted soils

The population size of the soil rhizobia in pots were assessed at sowing, 10 d after planting and at 30 d after planting. Approximately 2 g of soil was extracted from 2 pots in each treatment from between 5 to 10 cm below the pot surface using a sterile straw core. The soil extraction was approximately 5 cm away from the nearest plant. The exact weight of the samples was then determined before viable cell counts by the Miles and Misra plate count on TY containing tetracycline, with the initial 10-fold dilutions done on a weight to volume basis.

5.2.4 Data analysis

Data were analysed by means with standard errors, Pearson's correlation and where applicable, an analysis of variance (ANOVA) using SPSS version 22 (IBM Corp, released 2013). ANOVA was preceded by a test for normality and equal variances (Levene's test). Tukey's HSD was then used when ANOVA was found to be significant (statistical significance was set at 0.05). Competitiveness Index was calculated as per Section 4.2.2.3.

5.3 Results

5.3.1 Experiment 1: Rhizosphere competence of strains

To assess the role of rhizosphere competence in nodule occupancy and strain competition, the four rhizobial genotypes were co-inoculated with CIAT 899-GE at low densities (approximately 1,900 rhizobial cells seed^{-1}) and rhizosphere counts and nodule occupancies determined (Table 5.2).

When CIAT 899C and CIAT 899-GE were co-inoculated at an equal ratio, the ratio was maintained in the rhizosphere and in the nodules (Table 5.2). The NAK 287C/CIAT 899-GE co-inoculation resulted in a decline in the proportion of NAK 287C from 44% at inoculation to 14% in the rhizosphere as determined on tetracycline plates after 7 d. Unstained nodules were observed on plants in this treatment (Figure 5.2:C) and as these were attributabed to curing of pGM01 from NAK 287C (no nodules were found on uninoculated plants or unstained nodules in any other treatments), all unstained nodules were regarded as occupied by NAK 287C. Consequently, although NAK 287C appeared to multiply poorly in the rhizosphere (14%) in comparison to CIAT 899-GE (86%), 50% of the nodules were occupied by NAK 287C. However, the true density of NAK 287C in the rhizosphere of plants may have been higher than recorded because cells cured of pGM01 could not be enumerated on tetracycline plates.

In the NAK 210C/CIAT 899-GE treatment, 43% of the cells in the inoculum were NAK 210C but this proportion fell to 2% in the rhizosphere, resulting in 5% nodule occupancy (Table 5.2). In this treatment, greater rhizosphere multiplication of CIAT 899-GE resulted in a superior nodule occupancy. Lastly, the proportion of NAK 120C in the NAK 120C/CIAT 899-GE treatment declined from 41% at inoculation to 14% in the rhizosphere. Surprisingly, the strain did not occupy any nodules, suggesting that in addition to rhizosphere competence, other factors may influence nodule occupancy (or competitiveness) of a strain.

Table 5.2: Measured parameters following inoculation of CIAT 899-GE separately with CIAT 899C, NAK 120C, NAK 210C and NAK 287C

<table>
<tr><th>†STRAINS</th><th colspan="2">INOCULATION</th><th colspan="2">RHIZOSPHERE</th><th colspan="3">NODULATION</th></tr>
<tr><th></th><th>Average Inoculum numbers (cells seed^{-1})</th><th>Proportion (%)</th><th>Average Rhizosphere numbers (g^{-1} of matter‡)</th><th>Proportion (%)</th><th>Average Nodule numbers</th><th>Proportion (%)</th><th>Mean total nodules/root</th></tr>
<tr><td>899C|899-GE</td><td>824±*11|1049±17</td><td>44|56</td><td>$5.97\times10^{6}\pm4\times10^{5}$|$7.53\times10^{6}\pm1\times10^{6}$</td><td>45|55</td><td>65.8±9|83.8±16</td><td>44|56</td><td>149.6±25</td></tr>
<tr><td>NAK 287C|CIAT 899-GE</td><td>837±9|1063±6</td><td>44|56</td><td>$1.82\times10^{6}\pm6\times10^{5}$|$1.17\times10^{7}\pm2\times10^{6}$</td><td>14|86</td><td>66.4±10¥|65.6±11</td><td>50|50</td><td>132±21</td></tr>
<tr><td>NAK 210C|CIAT 899-GE</td><td>857±55|1137±61</td><td>43|57</td><td>$1.80\times10^{5}\pm5\times10^{4}$|$1.20\times10^{7}\pm2\times10^{6}$</td><td>2|98</td><td>4±2|82.5±15</td><td>5|95</td><td>86.5±17</td></tr>
<tr><td>NAK 120C|CIAT 899-GE</td><td>760±21|1073±60</td><td>41|59</td><td>$2.01\times10^{6}\pm6\times10^{5}$|$1.28\times10^{7}\pm3\times10^{6}$</td><td>14|86</td><td>0|87.5±29</td><td>0|100</td><td>87.5±29</td></tr>
</table>

† Strains co-inoculated together, separated by a bar. Subsequent data is also similarly separated to correspond to relevant strain

*Values after ± sign indicates standard error of means

‡Rhizosphere numbers are given per g of rhizosphere matter (roots and adhering soil)

¥unstained nodules were counted as occupied by NAK 287C as they were inferred to arise from infection by NAK 287C cells that had lost pGM01

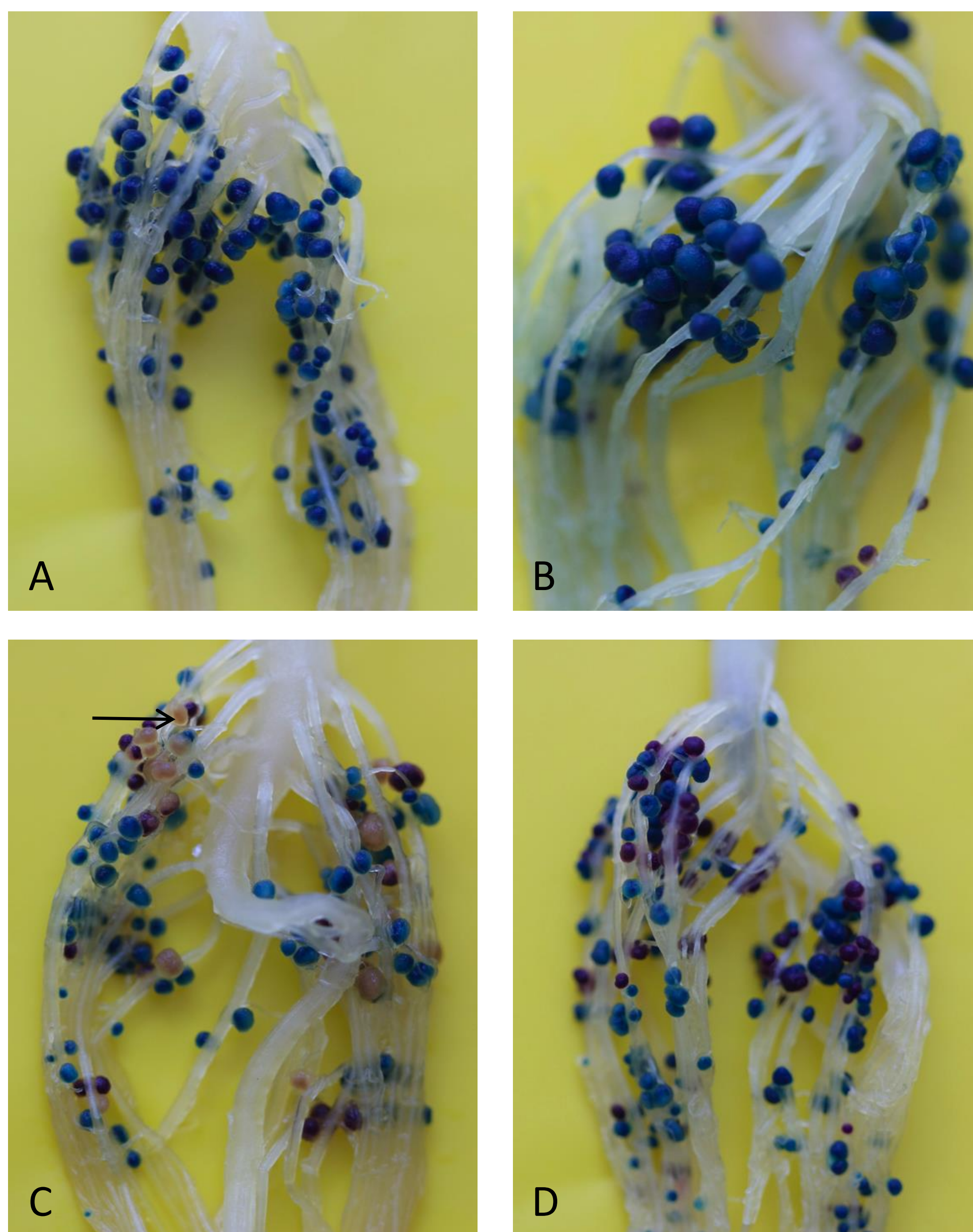


Figure 5.2: Photographs of stained *P vulgaris* cv. Kenya Tamu root nodules 17 d after inoculation with (**A**) NAK 120C/CIAT 899-GE (**B**) NAK 210C/CIAT 899-GE (**C**) NAK 287C/CIAT 899-GE and (**D**) CIAT 899C/CIAT 899-GE. Unstained nodules (arrow) were seen in the NAK 287C/CIAT 899-GE treatment (**C**).

5.3.2 Experiment 2: Competition between seed-inoculated CIAT 899-GE and soil-borne rhizobia

5.3.2.1 Rhizobial population densities in inoculated soil

Low cell density inoculation

Population estimates were not obtained from samples taken 2-4 h after soil inoculation as cell numbers were below the detection threshold for the enumeration method used. At planting, the populations of the four strains ranged between 4.4×10^4 and 8.7×10^4 cells g^{-1} of soil and did not differ significantly ($p<0.05$) from each other (Figure 5.3 & Appendix 5). Ten days later, the bacterial populations in the soils had increased, with an overall mean of $2.7 \times 10^5 \pm 6 \times 10^4$ cells g^{-1} of soil. At this sampling, homogenous soil population sizes (Tukey's HSD, P=0.05) were observed in all soils except soil inoculated with NAK 210C which had a population significantly lower than in NAK 287C and NAK 120C treatments (Figure 5.3 & Appendix 2). Except for NAK 287C treatment, the soil populations increased between day 10 and day 30 after planting. At 30 days, populations of all strain treatments differed significantly ($p<0.05$) from each other (Figure 5.3 & Appendix 5).

NAK 120C and NAK 210C treatments receiving supplementary N had similar population levels at planting and 10 d to those receiving no N. At 30 days, significant differences ($p<0.05$) were seen between the two N levels with the rhizobial populations being lower in the soil receiving supplementary N (Figure 5.3 & Appendix 5).

High cell density inoculation

The mean population of rhizobia, 2-4 h after inoculation was $1.2 \times 10^5 \pm 2 \times 10^4$ cells g^{-1} soil while at planting, the mean soil populations across the strain treatments ranged between 8.3×10^5 and 3.4×10^6 cells g^{-1} of soil. NAK 210C had the lowest population at this time but was only statistically ($p<0.05$) less than NAK 287C treatment (Figure 5.3 & Appendix 5). Ten days after planting, rhizobial populations in the pots had increased more than two-fold from those at planting to between 1.15×10^6 and 6.28×10^6 cells g^{-1} of soil. Statistically, the populations of NAK 120C and NAK 210C were different from those of NAK 287C and CIAT 899C. The period between day 10 and day 30 after planting saw slight population increases in NAK 120C and CIAT 899C and decreases in NAK 287C and NAK 210C. NAK 287C treatments suffered the most significant population decline during this period. At day 30, as was seen in the pots inoculated at

low density, the populations in all treatments differed statistically from each other (Figure 5.3 & Appendix 5).

The addition of nitrogen to NAK 120C and 120C treatments did not alter ($p<0.05$) population outcomes at the time of inoculation or by the day of planting in comparison to no nitrogen. However, from the 10th day after planting, strain-dependent population differences started to emerge (Figure 5.3 & Appendix 5).

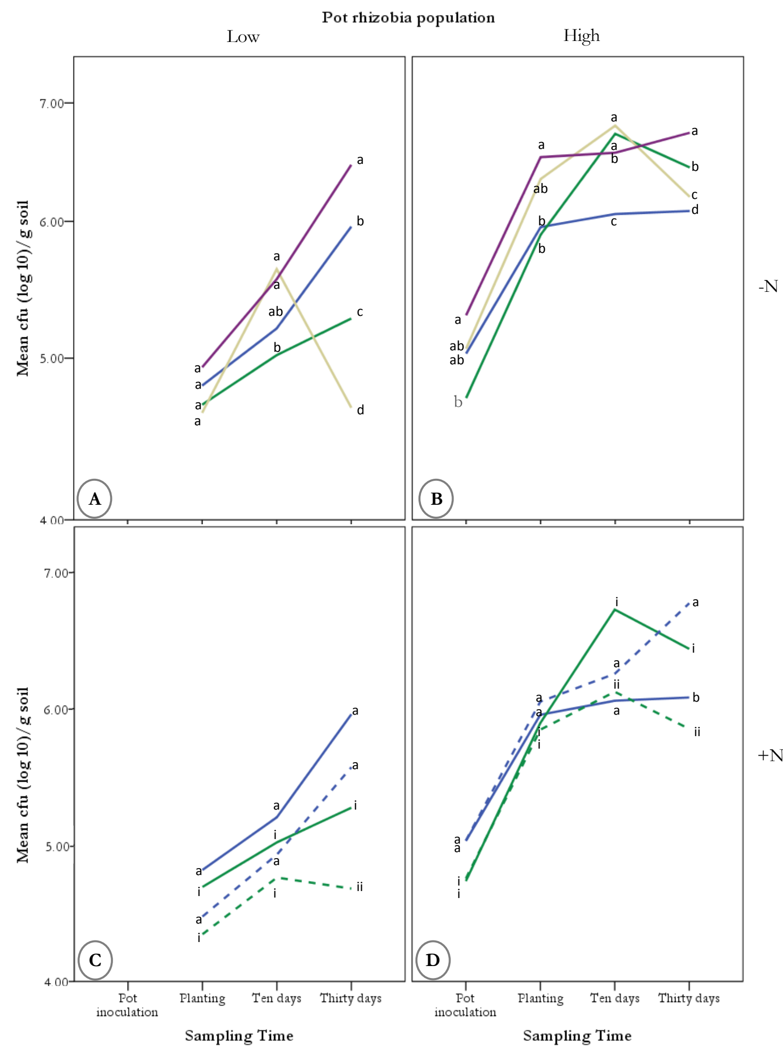


Figure 5.3: The rhizobia population (logarithmic cfu g^{-1} of soil) changes in the bulk soil inoculated separately with four strains (NAK 120C−/− −, NAK 210C−/− −, NAK 287C− and CIAT 899C−) at low (**A** and **C**) and high (**B** and **D**) cell densities over the experimental period. Panels **C** and **D** show the effect of supplementary N on soil populations of NAK 120C and NAK 210C following low and high cell density inoculations (dashed lines represent treatments with supplementary N). The low and high cell density soils were inoculated at 10^2 and 10^5 cells g^{-1} of soil respectively. Data points are means of duplicates and marked with the same letter, are not significantly different at 0.05 level using Tukey's HSD. For panels **C** and **D**, statistical comparisons are only made between the two N levels of the same strain treatment.

5.3.2.2 Nodule occupancies by seed-inoculated CIAT 899-GE

Strain differences in their ability to outcompete seed-inoculated CIAT 899-GE

When *P. vulgaris* inoculated with CIAT 899-GE was sown into soil containing NAK 120C (at either density), 65%-91% of the nodules were occupied by CIAT 899-GE, depending on other factors. The nodule occupancy by CIAT 899-GE ranged from 2%-15% in soils with NAK 210C, 1%-2% in soils with NAK 287C, and 1%-9% in soils with CIAT 899C (Table 5.3). An analysis of the means of the respective nodule occupancies revealed that the ability of strains to preclude CIAT 899-GE from nodules varied in increasing order from NAK 120C, NAK 210C, CIAT 899C to NAK 287C (Table 5.3) indicating that nodule occupancy by CIAT 899-GE is dependent on the rhizobial genotype in the soil.

Effect of rhizobial soil population density on nodule occupancy by seed-inoculated CIAT 899-GE

Seed-applied CIAT 899-GE formed 72%-90% of the nodules in soils containing NAK 120C at the low density (approx. 10^4 cells g^{-1} of soil at planting) and 65%-79% of the nodules in the soils with NAK 120C at the high density (approx. 10^6 cells g^{-1} of soil at planting) (Table 5.3). As the differences were significant ($P<0.05$), the data shows that the nodule occupancy by CIAT 899-GE was influenced by the density of NAK 120C in the soil, with a greater nodule occupancy by the inoculant at the lower soil rhizobial density. Similarly, nodule occupancy by CIAT 899-GE in soils carrying NAK 210C varied ($P<0.05$) depending on the density of rhizobia in the soil, with a 9%-15% occupancy in the low-density soil and 2%-5% in the high-density soil (Table 5.3).

In soils containing NAK 287C and CIAT 899C, nodule occupancy by CIAT 899-GE was greater in soil with the lower rhizobial density, in comparison to soil with higher rhizobial density, but the differences were not significant ($P>0.05$) (Table 5.3). This data indicates that while reducing the population density from 10^6 to 10^4 cells g^{-1} of soil at planting increased nodule occupancy by the inoculant, the magnitude of the effect was strain dependent.

Effect of inoculation rate on inoculant nodule occupancy

Two peat inoculants of CIAT 899-GE carrying 1.62×10^7 and 1.45×10^9 cells $gram^{-1}$ were prepared and used to inoculate *P. vulgaris* resulting in a low and high seed inoculation rate of $7.37 \times 10^4 \pm 3.8 \times 10^3$ and $6.58 \times 10^6 \pm 3.7 \times 10^5$ cells $seed^{-1}$ respectively. The low inoculation rate resulted in 72%-86% nodule occupancy by CIAT 899-GE in soils with NAK 120C, which was a significantly lower ($P<0.05$) than the 90%-91% nodule occupancy achieved with the higher inoculation rate (Table 5.3). Similarly, the higher inoculation rate improved nodule occupancy by the inoculant strain in soils carrying CIAT 899C from 1% to 5%-9%. In soils with NAK 210C and NAK 287C, varying the inoculation rate of CIAT 899-GE did not significantly ($P=0.05$) alter nodule occupancy (Table 5.3), indicating benefits from increasing inoculation rate are dependent on the genotype of rhizobia resident in the soil.

Effect of soil nitrogen on nodule occupancy by CIAT 899-GE

Only soils carrying NAK 120C and NAK 210C were subjected to variable N application. Nodulation was observed in all plants receiving supplementary N, indicating the amount of N applied was not sufficient to inhibit nodule formation, although nodules were smaller than those on plants receiving no supplementary N.

In both treatments (NAK 120C and NAK 210C), the addition of N had no effect ($P>0.05$) on the nodule occupancy outcomes of CIAT 899-GE (Table 5.3), suggesting changes in the soil N level have no consequences for nodule occupancy with these strains.

Table 5.3: Nodule occupancies attained by the seed-applied CIAT 899-GE and soil strains under the different experimental conditions

	Rhizobial genotype in soil	Rhizobial density at planting (g^{-1} soil)*	Inoculant rate (CIAT 899-GE $seed^{-1}$)**	Soil N***	Nodule occupancy by CIAT 899-GE (%) n=8	Nodule occupancy by soil strain (%) n=8
1	**NAK 120C**	6.7×10^4	L	0	**72**±6	**28**±6
2	"	"	H	0	**91**±3	**9**±3
3	"	3.1×10^4	L	1	**86**±3	**14**±3
4	"	"	H	1	**90**±3	**10**±3
5	"	0.91×10^6	L	0	**65**±5	**35**±5
6	"	"	H	0	**80**±3	**20**±3
7	"	1.1×10^6	L	1	**66**±8	**34**±8
8	"	"	H	1	**78**±6	**22**±6
9	**NAK 210C**	5.0×10^4	L	0	**15**±2	**85**±2
10	"	"	H	0	**14**±4	**86**±4
11	"	2.2×10^4	L	1	**10**±3	**90**±3
12	"	"	H	1	**9**±3	**91**±3
13	"	0.83×10^6	L	0	**3**±1	**97**±1
14	"	"	H	0	**4**±1	**96**±1
15	"	0.72×10^6	L	1	**2**±2	**98**±2
16	"	"	H	1	**5**±1	**95**±1
17	**NAK 287C**	4.4×10^4	L	0	**2**±1	**98**±1
18	"	"	H	0	**2**±.4	**98**±.4
19	"	2.2×10^6	L	0	**1**±.3	**99**±.3
20	"	"	H	0	**1**±.8	**99**±.8
21	**CIAT 899C**	0.87×10^5	L	0	**1**±.4	**99**±.4
22	"	"	H	0	**9**±2	**91**±2
23	"	3.4×10^6	L	0	**1**±.6	**99**±.6
24	"	"	H	0	**5**±1	**95**±1
25	0	0	0	0	**0**	**0**
26	0	0	0	2*	**0**	**0**

*SE are indicated in Appendix 5

**L=7.4×10^4, H=6.6×10^6

***0=no nitrogen, 1=1.2 mL of 1 M ammonium nitrate pot^{-1} $week^{-1}$, 2=2.4 mL of 1 M ammonium nitrate pot^{-1} $week^{-1}$

5.3.2.3 Ranking of factors

The effect sizes of the four factors on nodule occupancy outcomes were compared and partial eta squared measures calculated (Table 5.4). The genotype of the strain in the soil accounted for the biggest variance in nodule occupancy outcomes. The population density of the soil strain had the second highest effect on outcomes, and inoculum density was third. Nitrogen had the least effect on nodule occupancy by CIAT 899-GE. A significant interaction was only seen between rhizobial genotype and inoculum density (Table 5.4).

Table 5.4: The effect sizes of soil rhizobial genotype, background rhizobial population density, seed inoculation rate and nitrogen on nodule occupancy outcomes by CIAT 899-GE in *P. vulgaris* cv. Kenya Tamu.

Source	df	F	Sig.	Partial Eta Squared
Rhizobial genotype	3	800.533	.000	.935
Population density	1	24.232	.000	.126
Inoculum density	1	10.054	.002	.056
Nitrogen	1	.010	.922	.000
Genotype * Nitrogen	1	2.878	.092	.017
Genotype * Population density	3	2.492	.062	.043
Genotype * Inoculum density	3	6.388	.000	.102
Nitrogen * Population density	1	.103	.749	.001
Nitrogen * Inoculum density	1	1.586	.210	.009
Population density * Inoculum density	1	.473	.493	.003
Genotype * Nitrogen * Population density	1	3.433	.066	.020
Genotype * Nitrogen * Inoculum density	1	2.226	.138	.013
Genotype * Population density * Inoculum density	3	.092	.965	.002
Nitrogen * Population density * Inoculum density	1	1.014	.315	.006
Genotype * Nitrogen * Population density * Inoculum density	1	.620	.432	.004
Error	168			
Total	192			
Corrected Total	191			

5.4 Discussion

5.4.1 Soil rhizobial genotype as the main determinant of nodule occupancy by the inoculant

The results from this chapter indicated that the outcomes of inoculating *P. vulgaris* were mainly determined by the rhizobial genotype in the soil. Other factors - soil rhizobial density, inoculum rate, and soil N - had minimal (or no) effect on the nodule occupancy by the inoculant strain. This finding differs from that of Thies et al. (1991b), who reported that the success of inoculating legumes almost entirely depended on rhizobial densities in the soil. In the study by Thies et al. (1991b), the genotypes of the soil rhizobia were not surveyed.

Nodule occupancy data from Experiment 2 indicated that depending on the genotype of the soil strain, seed-inoculated CIAT 899-GE occupied between 1% and 91% of the nodules on *P. vulgaris* (Table 5.3). The genotype effects on nodule occupancies were significant ($P<0.05$) and a partial eta^2 of 0.935 indicated a large proportion of the variability in nodule occupancies was attributable to the rhizobial genotype (Table 5.4).

Results showed that although rhizobial density is important (partial eta^2 of 0.126), its effect is much lower than that of rhizobial genotype (Table 5.4). *P. vulgaris* is often cultivated in soils with high rhizobial densities in the range of 10^4 to 10^6 rhizobia g^{-1} soil (Alberton et al., 2006; Andrade et al., 1999; Anyango et al., 1995; Hungria et al., 2000; Kawaka et al., 2014; Langwerden, 2014; Thies et al., 1991b) and two densities of approximately 10^4 and 10^6 rhizobia g^{-1} soil at planting were established in a controlled experiment (Figure 5.3) to mimic field densities. Seed-applied CIAT 899-GE occupied a higher number of nodules in soils carrying *R. paranaense* NAK 120C and *Rhizobium* sp. NAK 210C at 10^4 cells g^{-1} soil in comparison to when soils carried 10^6 cells of either strain g^{-1} soil. In soils carrying *R. phaseoli* NAK 287C and CIAT 899C, soil densities did not alter nodule occupancy by the inoculant. NAK 120C and NAK 210C were the least competitive among the four strains used (Figure 5.1) indicating that population densities were only slightly important in soils with uncompetitive genotypes. In soils carrying the equally competitive CIAT 899C or the more competitive NAK 287C, nodule occupancy by CIAT 899-GE was the same at either soil density (Table 5.3).

Previous studies evaluating the effect of population sizes of indigenous rhizobia on inoculation of legumes used field sites characterised by MPN (Singleton & Tavares,

1986; Thies et al., 1991b), overlooking the roles played by the different rhizobial genotypes in the soil in influencing nodule occupancy by the inoculant. Consequently, different inoculation responses are reported in soils with similar rhizobial densities or, a low rhizobial density inhibits inoculation response in *P. vulgaris* (Thies et al., 1991a) while a higher one does not (Hungria et al., 2000; Hungria et al., 2003; Vlassak et al., 1996). These studies did not evaluate whether strain genotypes were present that either inhibited or allowed inoculation response.

In the current study, individual rhizobial genotypes were analysed, and in this deconstructed factorial, (which still allowed the experimental soil strain the positional advantage enjoyed by strains in field settings), the key role of rhizobial genotypes in inoculant nodule occupancy was examined. From the nodule occupancy data, inoculation responses would be expected in soils containing NAK 120C at up to 10^6 cells g^{-1} of soil while no response would be expected in soils containing NAK 287C at 10^4 cells g^{-1} of soil. Although it is unknown whether substantially higher inoculant nodule occupancies would have been achieved at lower densities of NAK 287C (e.g. at 100 cells g^{-1} of soil), such low densities are irrelevant to studies with *P. vulgaris* because cultivation of the crop is generally in soils with rhizobia in the range of 10^4-10^6 cells g^{-1} soil (Alberton et al., 2006; Andrade et al., 1999; Anyango et al., 1995; Hungria et al., 2000; Kawaka et al., 2014; Langwerden, 2014). Are soil rhizobial numbers important determinants of inoculation response in *P. vulgaris*? From the findings of the current study, the answer appears to be 'numbers of what?'.

Inoculation rates had a small effect on nodule occupancy as indicated by a partial eta^2 of 0.056 (Table 5.4). The inoculum dosages applied, of approximately 7×10^4 and 7×10^6 rhizobia per seed were, respectively, below and above the 10^5 rhizobia per seed recommended for the inoculation of *P. vulgaris* (Bullard et al., 2005; Hungria et al., 2005; Lupwayi et al., 2000). However, the findings here indicated that, in soils with rhizobia at densities found in soils used to cultivate *P. vulgaris*, a high inoculum density improved nodule occupancy only slightly, depending on the competing soil rhizobial genotype. The interaction effects of the inoculum dosage and rhizobial genotype were significant (Table 5.4), highlighting the importance of the genotype of the soil rhizobia in the final nodule occupancy outcomes of an inoculated legume.

The current study also found that a low level of supplementary mineral N had no effect on the nodule occupancy of the strains (Table 5.3; Table 5.4). These findings are

consistent with those of Abaidoo et al. (1990) but appear inconsistent with those of Caballero-Mellado and Martinez-Romero (1999) who reported that some rhizobial groups failed to nodulate *P. vulgaris* with the addition of mineral N. The N applied in the current study reduced nodule size. This rate may have been higher than found in soils used to cultivate *P. vulgaris,* especially in SSA, where nodule sizes have not been reported to be diminished by soil N. CIAT 899, NAK 120 and NAK 210 did not appear to be sensitive to mineral N as their nodule occupancies were unaffected by added N. Although N was found to not affect the nodule occupancy of the strains used, in view of reports of nitrate sensitivity with some strains (Caballero-Mellado & Martinez-Romero, 1999), screening of candidate inoculant strains for nitrate tolerance may be useful, because *P. vulgaris* is cultivated in soils heterogeneous in N. Such experiments would add an extra level of complexity to glasshouse experiments, but would make them more closely matched to field conditions.

5.4.2 Strain influence on nodule occupancy is effected through rhizosphere competence and preferential nodulation

a) Rhizosphere competence

Experiment 1 revealed that CIAT 899C, NAK 120C, NAK 210C and NAK 287C differed in their ability to colonise the rhizosphere of *P. vulgaris* when co-inoculated with CIAT 899-GE. In general, data revealed that strains exhibiting greater rhizosphere competence had greater success at occupying root nodules. Correspondingly, strains that poorly colonized the rhizospheres occupied fewer nodules. For example, CIAT 899-GE, which formed more nodules when co-inoculated with NAK 120C and NAK 210C at almost equal ratios, greatly outnumbered the two strains in the rhizosphere following inoculations (Table 5.2). The data from co-inoculation of NAK 287C and CIAT 899-GE appeared to be an exception to this rule, but rhizosperic numbers of NAK 287C were not reliably enumerated on tetracycline plates due to loss of pGM01 from this strain. For all other combinations, superior nodule occupancy by CIAT 899-GE was preceded by its superior rhizosphere numbers.

Rhizosphere competence is related to a range of factors including survival and growth of strains under prevalent environmental conditions such as pH (Anyango et al., 1998), as well as other characteristics intrinsic to strains such as growth rates (Li & Alexander, 1986), tolerance to microbial antagonisms (Mrabet et al., 2006), chemotaxis and motility

(Cooper, 2007), root attachment and biofilm formation (Janczarek et al., 2015a). Additionally, roots of legumes secrete flavonoids and other low molecular weight organic compounds such as sugars and amino acids that may favour the growth and multiplication of certain rhizobia. Streit et al. (1992) observed that the ability of strains to degrade aromatic compounds commonly secreted by plant roots was correlated with greater rhizobial competitiveness for nodulation of *P. vulgaris*. Although the basis for differences in rhizosphere colonization among the study strains were not investigated, rhizosphere competence was observed to be a key characteristic that determined nodule occupancy by the different rhizobial genotypes.

b) Preferential nodulation

Strains also differed in their relative 'affinities' for *P. vulgaris*. This difference was deduced from the observation that rhizospheric ratios of co-inoculated strains were (in instances) disproportionate to their nodule occupancy ratios. For example, 14% of rhizobia in the rhizosphere of plants inoculated with a suspension of NAK 120C/CIAT 899-GE were NAK 120C, but all nodules were occupied by CIAT 899-GE (Table 5.2). In experiment 2, the soil densities of the four strains differed only slightly for the period between planting and 10 days post planting when most nodulation events occur (George et al., 1992). However, despite having similar bulk soil populations the strains resulted in significantly different nodule occupancies (Table 5.3 & Appendix 5). Furthermore, the least competitive soil-borne strain in this experiment, NAK 120C, did not have the lowest cell density in the bulk soil (Figure 5.3 & Appendix 5). In both experiments, NAK 210C had the lowest rhizospheric numbers. But in both experiments, NAK 210C occupied a greater proportion of nodules against CIAT 899-GE than NAK 120C, indicating that nodule occupancy was not an entirely numerical phenomenon.

Is the observation that NAK 120C consistently occupied less nodules than NAK 210C (despite NAK 120C being more rhizosphere competent) evidence for preferential nodulation? Preferential nodulation, where a legume host nodulates with a strain selected from a pool of compatible strains, in disregard of population proportions, has previously been reported in *Vicia faba* (Laguerre et al., 2003), *P. vulgaris* (Montealegre & Graham, 1996; Montealegre et al., 1995) and *Trifolium* spp. (Yates et al., 2008) symbioses. In the study by Montealegre et al. (1995), *P. vulgaris* cv. RAB39 preferentially nodulated with *R. tropici* CIAT 899 in the presence of seven *R. etli* strains. While the

study by Montealegre et al. (1995) and others above described strains overcoming numerical disadvantages, the authors did not perform rhizobial counts following inoculation. It is, therefore, arguable that the population ratios may have altered substantially following co-inoculation, leading to results that relate to rhizosphere supremacy and not preferential nodulation. However, in the current study, NAK 120C consistently formed fewer nodules than would be predicted by its numerical proportion in the rhizosphere, assessed at the critical nodule forming periods. A logical explanation to this observation would be that the host was preferentially nodulating with CIAT 899-GE over NAK 120C.

While the basis for preferential nodulation is unknown, preferential nodulation is likely to result from incompatibilities at any of the many stages of a legume-rhizobia interaction (Section 1.2). It is highly plausible that a suboptimal interaction during the early stages of nodulation, between *P. vulgaris* and (for example) NAK 120C, led to nodule occupancies in favour of CIAT 899-GE. For example, greater attachment to root infection sites, mediated by EPS, can confer a competitive advantage to a strain (Janczarek et al., 2015b; Williams et al., 2008). Additionally, rhizobia are known to use EPS and effector proteins to evade host defences. Rhizobia that lack secretion systems for movement of effector proteins are still able to nodulate legumes (Deakin & Broughton, 2009) but the infection efficiency of these strains may be lower. Future studies should explore some of these possible bases of preferential nodulation using strains identified here.

5.4.3 Concluding remarks

The success of inoculating *P. vulgaris* in the field has often varied between sites (Graham & Ranalli, 1997). While population sizes of compatible rhizobia have in the past been linked to the variability of inoculation outcomes (Thies et al., 1991a), this study found that the population densities were not necessarily the most important factor. The rhizobial genotype in the soil was found to have the greatest effect on the outcomes of inoculating *P. vulgaris*. Using four genotypes, a continuum was observed in regards to ability to inhibit or allow nodulation by an inoculant strain. Many other genotypes of similar competitiveness to the ones used in this chapter were identified from Kenyan soils (Figure 4.21) and future analyses of the large pool of strains, using controlled experiments with greater complexity and genomic tools, is anticipated to continue to shed light on the seemingly random outcomes of inoculating *P. vulgaris*.

CHAPTER 6
General discussion

6.0 Recap of aims

The studies presented in this thesis addressed four key objectives. The first of these was to identify and characterise rhizobia that nodulate *P. vulgaris* in Kenya. This had previously not been completed to such detail despite numerous studies showing large populations of *P. vulgaris*-rhizobia in Kenyan soils (Anyango et al., 1995; Kawaka et al., 2014; Muthini et al., 2014). The second objective was to assess the indigenous rhizobia for their ability to fix N_2 with a view to identifying potential inoculant strains pre-adapted to Kenyan soils. These strains would serve as alternatives to the currently used strains especially in areas where current inoculant strains such as CIAT 899 have been shown to be poorly adapted to edaphic conditions (Anyango et al., 1998). Whereas the finding of effective strains capable of nodulating *P. vulgaris* was important, it was considered equally important to understand factors linked to outcomes of applying these strains as inoculants. Rhizobial competition has previously been linked to poor inoculation outcomes in *P. vulgaris* (Thies et al., 1991a). To facilitate rhizobial competition studies, the third objective was to develop a marker gene approach for use in competition studies involving diverse and numerous rhizobial strains, a protocol for which was unavailable. The fourth objective was to identify key determinants of successful inoculation of *P. vulgaris* from among four factors viz; rhizobial genotype, soil population density, inoculation rate and soil N.

6.1 Diversity of rhizobia from Kenyan soils

Analyses of 197 strains from diverse ecological zones in central, western and coastal Kenya (Table 2.3) by RP01-PCR, PCR-RFLP, and by comparisons of 16S rRNA and *recA* sequences revealed the strains were genetically diverse. These strains belong to at least five species of *Rhizobium* viz., *R. sophoriradicis*, *R. phaseoli*, *R. leucaenae*, *R. paranaense* and *R. etli*. In addition to these five species the representatives of the dominant PCR-RFLP group 1, to which 65% of the strains belonged (Table 2.4), were not identified at the species level. These strains had *recA* sequences with less than 96.6% identity to corresponding sequences of current type strains in the genus *Rhizobium*. This divergence was deemed sufficient to designate the strains as putative new species in the genus *Rhizobium* (see Section 2.4.1). As the purpose of the current study was not to describe species, the relative phylogeny of the study strains was considered sufficiently assessed without the need for analyses of more genes.

The findings on the diversity, phylogeny and taxonomy of rhizobia that nodulate *P. vulgaris* in Kenya were especially revealing of the distribution, survival and evolution of *P. vulgaris* symbionts in an area of recent introduction (400-500 years ago) (Greenway, 1944). Anyango et al. (1995), broadly hypothesized that the rhizobia that nodulate *P. vulgari* in Kenyan soils were either indigenous to Kenyan soils or were introduced with the crop. This hypothesis can be subdivided and refined into five hypotheses that predict the genetic nature of the strains as follows:

i. The strains were introduced with the crop e.g. inadvertently on *P. vulgaris* seed and therefore belong to species that nodulate *P. vulgaris* in its centers of diversity (Mesoamerica and the Andes) or in other *P. vulgaris*-growing regions of the world that export significant quantities of *P. vulgaris* to Kenya.
ii. The strains were introduced after the crop as inoculants and therefore belong to *R. tropici* or *R. etli*, the species to which the two main strains used to inoculate *P. vulgaris* in Kenya belong.
iii. The strains are indigenous to Kenyan soils and belong to described rhizobial species that do not traditionally nodulate *P. vulgaris* but that harbor symbiotic genes corresponding to symbiovars known to nodulate *P. vulgaris*.
iv. The strains are indigenous to Kenyan soils and belong to undescribed rhizobial species that harbor symbiotic genes corresponding to symbiovars known to nodulate *P. vulgaris*.
v. The strains are indigenous to Kenyan soils and belong to undescribed rhizobial species and carry unique symbiotic genes that nevertheless allow successful nodulation of *P. vulgaris*.

In the studies described in this thesis, I presented evidence that supports the tenability of hypotheses one, two, three and four but found no evidence in support of hypothesis five. In support of hypothesis one, some study strains are *R. etli*, *R. phaseoli* or *R. paranaense*. These species nodulate *P. vulgaris* in the centers of bean diversity (e.g. Mexico and Argentina) and major bean-producing countries such as Brazil, Colombia, Ethiopia and U.S.A (Aguilar et al., 2004; Aserse et al., 2012; Dall'Agnol et al., 2014; Lopez-Guerrero et al., 2012; Martinez-Romero, 2003). It is conceivable that the three species were carried from these regions into Kenyan soils on bean seed (Mora et al., 2014; Pèrez-Ramirez et al., 1998) as Kenya has had significant bean market ties to some of these countries over the years (Greenway, 1944; Katungi et al., 2009).

R. tropici CIAT 899 is the primary *P. vulgaris* inoculant strain in Kenya but *R. etli* USDA 2667 is also occasionally used (Bala et al., 2011; Koinange, 2015). None of the isolates belonged to *R. tropici,* but strains belonging to *R. etli* were recovered. The *R. etli* strains recovered were not compared to USDA 2667 and therefore, they may or may not be evidence for hypothesis two. Nodule sampling was from farms with no known history of rhizobial inoculation, and the absence of CIAT 899 or USDA 2667 from among the strains recovered would not be surprising.

The recovery of strains of *R. leucaenae* and *R. sophoriradicis* from *P. vulgaris* nodules supports the third hypothesis. The two species are not considered specialist *P. vulgaris* nodulators. *R. leucaenae* strains can nodulate *P. vulgaris* (Ribeiro et al., 2012) while *R. sophoriradicis* (Jiao et al., 2015) was (before the current study) not known to nodulate *P. vulgaris* outside of laboratory conditions. However, *R. leucaenae* and *R. sophoriradicis*, as was observed in isolates from this thesis (Figure 2.5), carry broad host range symbiotic genes (sv. tropici) and narrow host range symbiotic genes (sv. phaseoli) respectively (Jiao et al., 2015; Ormeño-Orrillo et al., 2012; Ribeiro et al., 2012) that enable them to nodulate *P. vulgaris* as well as *Leucaena* spp. and *Sophora* spp. Species in the latter two genera are present in Kenya (Franzel et al., 2014; Leonard et al., 2015) and it may therefore be speculated that these *R. leucaenae* and *R. sophoriradicis* strains recovered also nodulate the alternative legumes.

In support of hypothesis four, 65% of the strains recovered belonged to putative new species. The strains harbored *nodC* sequences that placed them in symbiovar phaseoli, with 100% identities to, for example, *nodC* sequences of *R. vallis* (Figure 2.5). The identical *nodC* sequences with strains in symbiovar phaseoli may be indicative of genetic exchange with symbiovar phaseoli strains of other rhizobial species. Phylogenetic evidence exists for the transfer of symbiotic genes among members of the genus *Rhizobium* (Kumar et al., 2015) while *in vitro* self-transmissibility has been demonstrated (Brom et al., 2000; Rao et al., 1994). The strains in the potentially new species may also have arisen from chromosomal speciation accompanied by the inheritance of conserved symbiotic genes.

There was no evidence in support of hypothesis five. All isolates tested had *nodC* sequences 100% identical to corresponding genes of known strains in described rhizobial species. The exception was *R. paranaense* NAK 120 which was found to carry

nodC sequences divergent from those in symbiovar tropici. However, the strain was isolated from *Albizia* sp. and it is unknown whether the strain nodulates *P. vulgaris in situ*.

The above finding that a limited range of conserved symbiotic genes are carried by the numerous species that nodulate *P. vulgaris* was intriguing. Although the same has been reported by other researchers (Aserse et al., 2012; Faghire et al., 2012; Jiao et al., 2015; Rogel et al., 2011; Rouhrazi et al., 2016), it warrants further comment here as it may be of greater relevance to the *P. vulgaris* symbiosis than has been considered before.

What drives the phenomenon of highly conserved symbiotic genes in diverse chromosomal backgrounds? One of the theories of speciation that may account for the observation is the recurrent niche invasion model (Wiedenbeck & Cohan, 2011). In this speciation model, a lineage diversifies over time, but with recurrent loss and acquisition of niche-determining genes. In *Rhizobium*, the symbiotic genes are niche-determining as they confer nodulation abilities. The loss and acquisition of symbiotic genes, over time, would result in diverse genetic backgrounds with conserved symbiotic genes as was seen in the current study and others. Exchange of symbiotic genes among strains in the genus *Rhizobium* has been demonstrated by phylogenetic evidence and *in vitro* experiments (Kumar et al., 2015; Rao et al., 1994). It may therefore be hypothesized that the broad-host-range of *P. vulgaris,* is only limited by the transmissibility of symbiotic genes among rhizobia.

Of what consequence is the phenomenon? Evolution of *P. vulgaris* symbionts through acquisition of symbiotic genes by strains with diverse chromosomal backgrounds may be of considerable consequence to the effectiveness of the *P. vulgaris* symbiosis. This is because the transfer of symbiotic genes into new genomic backgrounds can lead to strains of partial effectiveness (Martínez et al., 1987; Nandasena et al., 2007). From the effectiveness experiments in Chapter 3, strains belonging to the putative new species (mostly harbouring *nodC* type γ-a) were less effective at N_2 fixation (Figure 3.8). Whether this reduced effectiveness is related to a recent acquisition of the symbiotic genes needs further investigation. The second consequence is that the new symbionts (of likely suboptimal fixation) also exacerbate the challenge of rhizobial competition that already compromises attempts to inoculate *P. vulgaris*.

6.2 Candidate inoculant strains for Kenya

Using data on genetic diversity and origin of the 197 study strains, 52 were selected and evaluated for N_2 fixation on Kenyan cultivars of *P. vulgaris* in glasshouse experiments. Strains were found to be variably effective on *P. vulgaris* (Figure 3.2) and eleven strains were as effective at N_2 fixation as CIAT 899, a leading inoculant strain for *P. vulgaris*. These 11 strains (NAK 407, NAK 458, NAK 354, NAK 327, NAK 227, NAK 214, NAK 104, NAK 288, NAK 239, NAK 157 and NAK 299) are now candidate inoculant strains and may represent well adapted and highly effective inoculants for *P. vulgaris* in Kenya. The candidate strains were isolated from ecologically diverse sites differing in soil types, rainfall, and temperatures (Table 2.2) and are therefore expected to be adapted to the environmental stresses present in those areas. These strains could be used as alternatives to CIAT 899 to inoculate *P. vulgaris* in ecological zones where the failure of inoculation with *R. tropici* CIAT 899 is believed to be due to poor adaptability or low saprophytic competence (Anyango et al., 1998; Gicharu et al., 2013). However, the high level of symbiotic performance by the strains observed under controlled conditions will need to be demonstrated in field evaluations. Furthermore, the strains will need to be assessed for genetic stability as, previously, some strains have been shown to lose their ability to nodulate over periods of use (Bullard et al., 2005).

6.3 Limitation of N_2 fixation in *P. vulgaris*

Although effective strains were identified, N_2 fixation in *P. vulgaris* was observed to be restricted in two ways. The first was through regulation of shoot N to within a range of 3.7%-4.1% N, outside of which plant biomass was greatly diminished (Figure 3.3). It is speculated here that this range allows the symbiotic *P. vulgaris* plant the optimal control of growth rate, which tends be co-limited by N and C (Andrews et al., 2007). However, this range of shoot N concentration resulted in approximately 60% less biomass than produced by plants grown with mineral N supplied at the maximum rate, meaning this regulation in *P. vulgaris* may limit the ability of N_2 fixation to stimulate growth comparable to that achieved under mineral N. Future studies exploring this finding would benefit from a greater number of rhizobial strains, *P. vulgaris* cultivars, and experimental set ups such as pot volumes. Inclusion of other legumes would also clarify whether similar shoot N-biomass relationships exist more widely in legumes.

Secondly, N_2-fixing *P. vulgaris* plants were observed to have a low N utilization efficiency in comparison to N-fed plants, whereby for a given tissue % N, they had considerably lower biomass (Figure 3.11) (see Section 3.4.5). These data suggest that N_2 fixation in *P. vulgaris* occurs at considerable energy costs to the host plant. Indeed, any N_2 fixation is energy intensive with the reduction of each mole of N_2 requiring eight moles of protons, eight moles of electrons, and 16 moles of ATP (Lodwig & Poole, 2010). Additionally, nitrate uptake, assimilation, protein synthesis and nodule organogenesis are processes that require energy. However, plants differ in their general efficiencies in N_2 fixation with estimates of 3.1 mg of C used per N fixed in cowpea and 6.6 mg of C per N fixed in lupin (Layzell et al., 1979). Calculations from data on Figure 3.11 indicate a loss of shoot C of 8.4 mg C per mg N fixed, which is slightly greater than the lupin value of 6.6 mg C per mg N fixed. This suggests that perhaps N-fixing *P. vulgaris* burns off C at a greater rate than other legumes. Differences in efficiencies of N_2 fixation occur from factors such as differences in nodulation patterns and carbon requirements of nitrogenase (Witty et al., 1983).

It is important to note that N_2 fixation experiments in this study measured vegetative yield and not grain yield. Grain yield was not measured because glasshouse experiments would not be expected to give meaningful grain yield due to the limited soil volume available to plants grown in pots. Vegetative yield is not perfectly correlated with grain yield (Masclaux-Daubresse et al., 2010) and therefore, future studies in this area should include field experiments that measure grain yield in order to obtain a clearer picture on the limitation of *P. vulgaris* symbiosis.

6.4 Broad-host-range plasmid for rhizobial competition assays

Inoculant strains for *P. vulgaris* often have to compete with strains resident in soils for nodule formation. Therefore, in addition to inoculant strains being effective, they also need to be competitive for nodulation. To aid in the study of rhizobial competition, work in this thesis leveraged on existing marker gene technology to develop a broad host range plasmid (pGM01) that expressed *celB*. This plasmid was stable in 80% of the strains tested and proved to be an effective method to rapidly mark diverse rhizobial strains for competition studies.

The broad-host-range pGM01 (Figure 4.12) carried *celB*, a robust gene marker previously used in rhizobial competition studies via mini-transposons (Sanchez-

Canizares & Palacios, 2013; Sessitsch et al., 1996) Tests confirmed that the pGM01 could be maintained stably, in the absence of antibiotic selection, both *in vitro* and *in planta,* in diverse rhizobial chromosomal backgrounds. Additionally, pGM01 was shown to have no measurable effects, in diverse rhizobial strains, on the growth rates of the transconjugants (Figure 4.16), indicating an insignificant metabolic burden on the host cells under the conditions tested.

In this study, the plasmid-borne marker was successfully paired with *gusA* on a mini-transposon to facilitate a rapid identification of nodule occupants in competition assays, and it is anticipated that pGM01 will, in the future, continue to be a useful tool in rhizobial competition assays under controlled laboratory conditions. The need for confinement of genetically modified organisms in most countries means that marked strains cannot be used in field experiments. This reduces the utility of the marker but fairly complex and informative experiments can still be conducted in confinement using strains marked with pGM01.

6.5 Determinants of nodule occupancy by the inoculant in *P. vulgaris*

Strains marked with *gusA* and *celB* were used to evaluate the effects of four factors on the nodule occupancy by CIAT 889 in *P. vulgaris*. Results revealed that the competing rhizobial genotype was the key determinant of nodule occupancy by an inoculant strain. The other two factors found to be marginally important to nodule occupancy by the inoculant (soil rhizobial density and inoculation density) had interactions with the rhizobial genotype (Table 5.4) further highlighting the importance of the soil rhizobial genotype to the nodule occupancy outcomes of inoculating *P. vulgaris.*

It is important to note that in experiments conducted, nodule occupancies were evaluated, and not response to inoculation *per se*. The latter is measured through assessment of plant growth and yield (Yates et al., 2016b). However, predictions can be made from previous reports that have correlated nodule occupancy with response to inoculation. Thies et al. (1991a) estimated that in the presence of ineffective strains, at least 66% nodule occupancy is required by the inoculant strain to elicit a response, although compensatory mechanisms such as the development of large nodules may allow a response with fewer effective nodules (Singleton & Stockinger, 1983), which explains inoculation responses in Brazil with inoculants occupying less than 50% of the nodules (Hungria et al., 2003). Substantial nodule occupancies (65%-91%) by the seed-

applied inoculant strain were achieved when the poorly effective NAK 120C was in the soil, and a response to inoculation under these circumstances was likely. A nodule occupancy of up to 15% was achieved by CIAT 899-GE in soils containing the ineffective NAK 210C but a response to inoculation would have been unlikely considering the high energy cost associated with the formation of such a high proportion of ineffective nodules. NAK 287C is an effective strain and no response would be expected to inoculation, with the slightly more effective CIAT 899-GE occupying ≤2% of the total nodules.

The inoculation response predictions indicate that in some instances, depending on the genotype combinations (e.g. with NAK 120 vs. CIAT 899), it might be possible to get an inoculation response in *P. vulgaris* in soils with high rhizobial densities. For these combinations, rhizosphere competence and preferential nodulation dictate nodule occupancy (Section 5.4.2). The prevalence of genotypes in nature that would interact similarly with the inoculant and the host is unknown but reports of inoculation response in soils with high rhizobial densities (Hungria et al., 2000; Vlassak et al., 1996) may occur from such interactions.

Interestingly, it was found that in the absence of preferential nodulation, even the uncompetitive strain NAK 210 hindered the competitive seed-applied CIAT 899 from occupying an appreciable proportion of nodules. When evaluated using liquid cultures, NAK 210C was highly uncompetitive, forming only 5% of the nodules (Table 5.2) but occupied 85%-98% of nodules when pre-established in the soil (Table 5.3) due to the positional advantage. Previous studies with soybean also highlighted the huge positional disadvantage suffered by seed-applied inoculants (López-García et al., 2002) and here this disadvantage was shown to be substantial even with uncompetitive strains.

Rhizobial genotypes such as NAK 120 that can easily be overcome through seed inoculation may not always be present in soils (or at least not as homogeneous populations) and, therefore, recommendations to increase inoculation response in *P. vulgaris* include the following. The first is that alternative approaches, such as the application of granular inoculant formulations at tillage, be tested with *P. vulgaris* to reduce the positional disadvantage faced by inoculants applied on seeds. Secondly, as population densities were also found to affect nodule occupancy, the lowering of rhizobial densities in the soil prior to cultivation of *P. vulgaris* might be a useful intervention. Rhizobial numbers in the soil are closely linked to the cropping history of

the legume (Hungria et al., 2003; Vlassak et al., 1996) and it is logically expected that not growing the crop for several seasons would reduce densities of compatible rhizobia to levels where inoculation may lead to responses. Future studies can use approaches developed in this thesis to further test lower population levels, of various genotypes, for inoculation responses.

In conclusion, the work done in this thesis contributed to the understanding of the *P. vulgaris* symbiosis in several ways. In addition to identifying the rhizobia that nodulate *P. vulgaris* in Kenyan soils, the study findings on the putative new species potentially add to the rhizobial species known to support symbiosis with *P. vulgaris*. *P. vulgaris* in Kenya was observed to symbiotically associate with different species of rhizobia that carry conserved nodulation genes, giving insights into the evolution of *P. vulgaris* symbionts in an area where the crop has only been recently introduced. Strains effective at N_2 fixation on Kenyan cultivars of *P. vulgaris* were identified from among the strains recovered and these may represent well-adapted inoculant strains for *P. vulgaris* in Kenya. The study revealed rhizobial genotypes are the key determinants of inoculation responses in *P. vulgaris,* which was contrary to previous studies that had overemphasized the role of soil rhizobial densities. It is anticipated that future studies will build on findings in this study, as well as use the gene marker tools and approaches developed in this study to generate knowledge that will substantially reduce the current uncertainty surrounding the inoculation of *P. vulgaris*.

Appendices

Appendix 1: Estimates of evolutionary divergence between 16S rRNA gene sequences of study strains and of closely related described species. The number of base substitutions per site from between sequences is shown. Analyses were conducted using the Tamura 3-parameter model. The rate variation among sites was modelled with a gamma distribution (shape parameter = 1). The analysis involved 45 nucleotide sequences. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. There were a total of 1291 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013).

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
1	NAK 103																																													
2	NAK 120	0.022																																												
3	NAK 157	0.002	0.022																																											
4	NAK 210	0.006	0.018	0.008																																										
5	NAK 239	0.000	0.022	0.002	0.006																																									
6	NAK 245	0.002	0.022	0.002	0.008	0.002																																								
7	NAK 266	0.006	0.018	0.008	0.000	0.006	0.008																																							
8	NAK 287	0.005	0.023	0.006	0.007	0.005	0.006	0.007																																						
9	NAK 294	0.023	0.006	0.024	0.020	0.023	0.023	0.020	0.024																																					
10	NAK 312	0.006	0.018	0.008	0.000	0.006	0.008	0.000	0.007	0.020																																				
11	NAK 334	0.006	0.018	0.008	0.000	0.006	0.008	0.000	0.007	0.020	0.000																																			
12	NAK 349	0.022	0.000	0.022	0.018	0.022	0.022	0.018	0.023	0.006	0.018	0.018																																		
13	NAK 351	0.004	0.022	0.003	0.010	0.004	0.002	0.010	0.009	0.024	0.010	0.010	0.022																																	
14	NAK 358	0.006	0.018	0.008	0.000	0.006	0.008	0.000	0.007	0.020	0.000	0.000	0.018	0.010																																
15	NAK 382	0.007	0.017	0.008	0.001	0.007	0.009	0.001	0.008	0.019	0.001	0.001	0.017	0.009	0.001																															
16	NAK 384	0.063	0.062	0.064	0.068	0.063	0.063	0.068	0.066	0.060	0.068	0.068	0.062	0.063	0.068	0.067																														
17	NAK 387	0.002	0.021	0.003	0.009	0.002	0.002	0.009	0.007	0.022	0.009	0.009	0.021	0.002	0.009	0.008	0.061																													
18	NAK 458	0.002	0.022	0.002	0.008	0.002	0.003	0.008	0.006	0.023	0.008	0.008	0.022	0.002	0.008	0.007	0.063	0.002																												
19	R. acidisoli FH13(T) (KJ921033)	0.006	0.018	0.008	0.000	0.006	0.008	0.000	0.007	0.020	0.000	0.000	0.018	0.010	0.000	0.001	0.068	0.009	0.008																											
20	R. anhuiense CCBAU 23252(T) (KF111868)	0.006	0.018	0.008	0.000	0.006	0.008	0.000	0.007	0.020	0.000	0.000	0.018	0.010	0.000	0.001	0.068	0.009	0.008	0.000																										
21	R. azibense 23C2(T) (JN624691)	0.024	0.032	0.026	0.031	0.024	0.025	0.031	0.029	0.032	0.031	0.031	0.032	0.027	0.031	0.031	0.061	0.026	0.026	0.031	0.031																									
22	R. calliandrae CCGE524(T) (JX855162)	0.021	0.004	0.022	0.017	0.021	0.021	0.017	0.022	0.003	0.017	0.017	0.004	0.022	0.017	0.017	0.061	0.020	0.021	0.017	0.017	0.030																								
23	R. etli CFN 42(T) (CP000133)	0.004	0.022	0.003	0.010	0.004	0.002	0.010	0.009	0.024	0.010	0.010	0.022	0.000	0.010	0.009	0.063	0.002	0.002	0.010	0.010	0.027	0.022																							
24	R. fabae CCBAU 33202(T) (DQ835306)	0.001	0.022	0.002	0.007	0.001	0.002	0.007	0.005	0.024	0.007	0.007	0.022	0.004	0.007	0.008	0.064	0.003	0.002	0.007	0.007	0.025	0.022	0.004																						
25	R. freirei PRF 81(T) (EU488742)	0.025	0.008	0.026	0.022	0.025	0.025	0.022	0.026	0.005	0.022	0.022	0.008	0.026	0.022	0.021	0.063	0.024	0.025	0.022	0.022	0.032	0.004	0.026	0.026																					
26	R. gallicum R602sp(T) (AF008130)	0.006	0.018	0.008	0.000	0.006	0.008	0.000	0.007	0.020	0.000	0.000	0.018	0.010	0.000	0.001	0.068	0.009	0.008	0.000	0.000	0.031	0.017	0.010	0.007	0.022																				
27	R. hainanense I66(T) (U71078)	0.027	0.008	0.027	0.023	0.027	0.027	0.023	0.027	0.006	0.023	0.023	0.008	0.027	0.023	0.022	0.064	0.026	0.027	0.023	0.023	0.034	0.005	0.027	0.027	0.002	0.023																			
28	R. jaguaris CCGE525(T) (JX855169)	0.023	0.002	0.024	0.020	0.023	0.023	0.020	0.024	0.005	0.020	0.020	0.002	0.024	0.020	0.019	0.061	0.022	0.023	0.020	0.020	0.034	0.005	0.024	0.024	0.008	0.020	0.008																		
29	R. laguerreae FB206(T) (JN558651)	0.006	0.018	0.008	0.000	0.006	0.008	0.000	0.007	0.020	0.000	0.000	0.018	0.010	0.000	0.001	0.068	0.009	0.008	0.000	0.000	0.031	0.017	0.010	0.007	0.022	0.000	0.023	0.020																	
30	R. leguminosarum USDA 2370(T) (U29386)	0.009	0.021	0.010	0.002	0.009	0.010	0.002	0.009	0.022	0.002	0.002	0.021	0.013	0.002	0.003	0.071	0.011	0.010	0.002	0.002	0.033	0.020	0.013	0.009	0.024	0.002	0.026	0.022	0.002																
31	R. leucaenae USDA 9039(T) (X67234)	0.023	0.006	0.024	0.020	0.023	0.023	0.020	0.024	0.000	0.020	0.020	0.006	0.024	0.020	0.019	0.060	0.022	0.023	0.020	0.020	0.032	0.003	0.024	0.024	0.005	0.020	0.006	0.005	0.020	0.022															
32	R. lusitanum P1-7(T) (AY738130)	0.022	0.004	0.023	0.019	0.022	0.022	0.019	0.023	0.003	0.019	0.019	0.004	0.023	0.019	0.018	0.059	0.022	0.022	0.019	0.019	0.034	0.004	0.023	0.023	0.006	0.019	0.006	0.002	0.019	0.022	0.003														
33	R. mayense CCGE526(T) (JX855172)	0.021	0.002	0.022	0.017	0.021	0.021	0.017	0.022	0.005	0.017	0.017	0.002	0.022	0.017	0.017	0.060	0.020	0.021	0.017	0.017	0.032	0.002	0.022	0.022	0.006	0.017	0.006	0.004	0.017	0.020	0.005	0.002													
34	R. mongolense ATCC BAA-116(T) (U89817)	0.020	0.037	0.022	0.026	0.020	0.020	0.026	0.025	0.041	0.026	0.026	0.037	0.022	0.026	0.027	0.064	0.020	0.022	0.026	0.026	0.014	0.038	0.022	0.021	0.041	0.026	0.043	0.039	0.026	0.027	0.041	0.040	0.038												
35	R. multihospitium CCBAU 83401(T) (EF035074)	0.025	0.008	0.026	0.022	0.025	0.025	0.022	0.026	0.005	0.022	0.022	0.008	0.026	0.022	0.021	0.063	0.024	0.025	0.022	0.022	0.032	0.004	0.026	0.026	0.000	0.022	0.002	0.008	0.022	0.024	0.005	0.006	0.006	0.041											
36	R. paranaense PRF 35(T) (EU488753)	0.022	0.000	0.022	0.018	0.022	0.022	0.018	0.023	0.006	0.018	0.018	0.000	0.022	0.018	0.017	0.062	0.021	0.022	0.018	0.018	0.032	0.004	0.022	0.022	0.008	0.018	0.008	0.002	0.018	0.021	0.006	0.004	0.002	0.037	0.008										
37	R. phaseoli ATCC 14482(T) (EF141340)	0.001	0.022	0.002	0.007	0.001	0.002	0.007	0.005	0.024	0.007	0.007	0.022	0.005	0.007	0.008	0.064	0.003	0.002	0.007	0.007	0.025	0.022	0.005	0.002	0.026	0.007	0.027	0.024	0.007	0.009	0.024	0.023	0.022	0.021	0.026	0.022									
38	R. pisi DSM 30132(T) (AY509899)	0.001	0.022	0.001	0.007	0.001	0.001	0.007	0.005	0.024	0.007	0.007	0.022	0.003	0.007	0.008	0.064	0.003	0.002	0.007	0.007	0.025	0.022	0.003	0.001	0.026	0.007	0.027	0.024	0.007	0.009	0.024	0.023	0.022	0.021	0.026	0.022	0.002								
39	R. pusense NRCPB10(T) (FJ969841)	0.062	0.062	0.063	0.067	0.062	0.062	0.067	0.065	0.060	0.067	0.067	0.062	0.062	0.067	0.066	0.001	0.060	0.062	0.067	0.067	0.060	0.062	0.062	0.063	0.064	0.067	0.065	0.062	0.067	0.070	0.060	0.061	0.061	0.064	0.064	0.062	0.063	0.063							
40	R. radiobacter ATCC 19358(T) (AJ389904)	0.055	0.055	0.056	0.060	0.055	0.055	0.060	0.058	0.052	0.060	0.060	0.055	0.054	0.060	0.059	0.008	0.052	0.055	0.060	0.060	0.053	0.054	0.054	0.056	0.057	0.060	0.058	0.054	0.060	0.062	0.052	0.052	0.052	0.057	0.057	0.055	0.056	0.056	0.007						
41	R. sophorae CCBAU 03386(T) (KJ831229)	0.006	0.018	0.008	0.000	0.006	0.008	0.000	0.007	0.020	0.000	0.000	0.018	0.010	0.000	0.001	0.068	0.009	0.008	0.000	0.000	0.031	0.017	0.010	0.007	0.022	0.000	0.023	0.020	0.000	0.002	0.020	0.019	0.017	0.026	0.022	0.018	0.007	0.007	0.067	0.060					
42	R. sophoriradicis CCBAU 03470(T) (KJ831225)	0.002	0.021	0.003	0.009	0.002	0.002	0.009	0.007	0.022	0.009	0.009	0.021	0.002	0.009	0.008	0.061	0.000	0.002	0.009	0.009	0.026	0.020	0.002	0.003	0.024	0.009	0.026	0.022	0.009	0.011	0.022	0.022	0.020	0.020	0.024	0.021	0.003	0.003	0.060	0.052	0.009				
43	R. tropici CIAT 899(T) (CPOO4015)	0.024	0.009	0.026	0.021	0.024	0.024	0.021	0.025	0.006	0.021	0.021	0.009	0.027	0.021	0.022	0.064	0.025	0.026	0.021	0.021	0.032	0.005	0.027	0.025	0.001	0.021	0.002	0.009	0.021	0.023	0.006	0.007	0.007	0.040	0.001	0.009	0.025	0.025	0.065	0.058	0.021	0.025			
44	R. vallis CCBAU 65647(T) (FJ839677)	0.017	0.005	0.017	0.024	0.017	0.016	0.024	0.022	0.012	0.024	0.024	0.005	0.017	0.024	0.023	0.059	0.017	0.017	0.024	0.024	0.028	0.009	0.017	0.017	0.013	0.024	0.013	0.007	0.024	0.027	0.012	0.009	0.008	0.033	0.013	0.005	0.018	0.017	0.060	0.052	0.024	0.017	0.014		
45	B. japonicum USDA 6(T) (AP012206)	0.119	0.134	0.119	0.126	0.119	0.120	0.126	0.124	0.137	0.126	0.126	0.134	0.120	0.126	0.125	0.139	0.119	0.119	0.126	0.126	0.128	0.136	0.120	0.118	0.132	0.126	0.132	0.133	0.126	0.130	0.137	0.133	0.133	0.122	0.132	0.134	0.120	0.119	0.141	0.137	0.126	0.119	0.134	0.127	

Appendix 2: Estimates of evolutionary divergence between *recA* sequences of study strains and of closely related described species. The number of base substitutions per site from between sequences is shown. Analyses were conducted using the Tamura 3-parameter model. The rate variation among sites was modelled with a gamma distribution (shape parameter = 1). The analysis involved 45 nucleotide sequences. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. There were a total of 342 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013).

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
1	NAK 103																																													
2	NAK 120	0.127																																												
3	NAK 157	0.103	0.106																																											
4	NAK 210	0.091	0.115	0.095																																										
5	NAK 239	0.000	0.127	0.103	0.091																																									
6	NAK 245	0.108	0.110	0.034	0.092	0.108																																								
7	NAK 266	0.091	0.115	0.095	0.000	0.091	0.092																																							
8	NAK 287	0.003	0.132	0.107	0.087	0.003	0.103	0.087																																						
9	NAK 294	0.141	0.041	0.132	0.133	0.141	0.133	0.133	0.137																																					
10	NAK 312	0.091	0.115	0.095	0.000	0.091	0.092	0.000	0.087	0.133																																				
11	NAK 334	0.126	0.151	0.114	0.054	0.126	0.111	0.054	0.122	0.170	0.054																																			
12	NAK 349	0.146	0.031	0.128	0.147	0.146	0.137	0.147	0.151	0.041	0.147	0.176																																		
13	NAK 351	0.091	0.115	0.095	0.000	0.091	0.092	0.000	0.087	0.133	0.000	0.054	0.147																																	
14	NAK 358	0.091	0.115	0.095	0.000	0.091	0.092	0.000	0.087	0.133	0.000	0.054	0.147	0.000																																
15	NAK 382	0.091	0.115	0.095	0.000	0.091	0.092	0.000	0.087	0.133	0.000	0.054	0.147	0.000	0.000																															
16	NAK 384	0.149	0.129	0.159	0.137	0.149	0.164	0.137	0.153	0.139	0.137	0.189	0.134	0.137	0.137	0.137																														
17	NAK 387	0.047	0.126	0.107	0.080	0.047	0.103	0.080	0.044	0.130	0.080	0.106	0.135	0.080	0.080	0.080	0.157																													
18	NAK 458	0.000	0.127	0.103	0.091	0.000	0.108	0.091	0.003	0.141	0.091	0.126	0.146	0.091	0.091	0.091	0.149	0.047																												
19	R. acidisoli FH13(T) (KJ921098)	0.101	0.135	0.098	0.034	0.101	0.103	0.034	0.097	0.144	0.034	0.064	0.158	0.034	0.034	0.034	0.149	0.082	0.101																											
20	R. anhuiense CCBAU 23252(T) (KF111980)	0.098	0.145	0.112	0.051	0.098	0.108	0.051	0.094	0.147	0.051	0.090	0.162	0.051	0.051	0.051	0.154	0.084	0.098	0.061																										
21	R. azibense 23C2(T) (JN624700)	0.146	0.116	0.143	0.134	0.146	0.138	0.134	0.146	0.121	0.134	0.173	0.125	0.134	0.134	0.134	0.144	0.146	0.146	0.151	0.144																									
22	R. calliandrae CCGE524(T) (JX855189)	0.142	0.082	0.143	0.135	0.142	0.139	0.135	0.138	0.078	0.135	0.184	0.086	0.135	0.135	0.135	0.143	0.150	0.142	0.146	0.135	0.113																								
23	R. etli CFN 42(T) (CP000133)	0.116	0.123	0.044	0.099	0.116	0.021	0.099	0.111	0.146	0.099	0.122	0.146	0.099	0.099	0.099	0.173	0.103	0.116	0.111	0.100	0.151	0.139																							
24	R. fabae CCBAU 33202(T) (EF5799941)	0.106	0.126	0.107	0.072	0.106	0.107	0.072	0.102	0.126	0.072	0.110	0.149	0.072	0.072	0.072	0.182	0.102	0.106	0.092	0.072	0.157	0.142	0.115																						
25	R. freirei PRF 81(T) (EU488827)	0.126	0.086	0.122	0.126	0.126	0.114	0.126	0.121	0.087	0.126	0.160	0.095	0.126	0.126	0.126	0.121	0.133	0.126	0.129	0.134	0.142	0.105	0.130	0.147																					
26	R. gallicum R602sp(T) (CPOO6877)	0.147	0.125	0.157	0.139	0.147	0.144	0.139	0.142	0.109	0.139	0.169	0.135	0.139	0.139	0.139	0.159	0.152	0.147	0.157	0.141	0.034	0.118	0.157	0.157	0.148																				
27	R. hainanens _I66(T) (HQ394252)	0.113	0.071	0.130	0.122	0.113	0.131	0.122	0.109	0.071	0.122	0.146	0.079	0.122	0.122	0.122	0.128	0.120	0.113	0.124	0.130	0.137	0.112	0.147	0.134	0.037	0.143																			
28	R. jaguaris CCGE525(T) (JX855192)	0.168	0.071	0.153	0.152	0.168	0.144	0.152	0.163	0.080	0.152	0.178	0.072	0.152	0.152	0.152	0.166	0.164	0.168	0.159	0.187	0.156	0.083	0.153	0.169	0.127	0.163	0.110																		
29	R. laguerreae FB206(T) (JN558681)	0.139	0.143	0.118	0.087	0.139	0.132	0.087	0.135	0.131	0.087	0.113	0.131	0.087	0.087	0.087	0.180	0.110	0.139	0.086	0.084	0.150	0.136	0.133	0.090	0.147	0.161	0.134	0.135																	
30	R. leguminosarum USDA 2370(T) (AJ294376)	0.106	0.122	0.103	0.051	0.106	0.099	0.051	0.102	0.119	0.051	0.082	0.132	0.051	0.051	0.051	0.160	0.087	0.106	0.057	0.063	0.122	0.118	0.104	0.076	0.125	0.132	0.121	0.148	0.037																
31	R. leucaenae USDA 9039(T) (AJ294372)	0.137	0.044	0.137	0.120	0.137	0.148	0.120	0.133	0.021	0.120	0.166	0.059	0.120	0.120	0.120	0.129	0.126	0.137	0.140	0.146	0.112	0.093	0.161	0.122	0.091	0.101	0.075	0.099	0.140	0.127															
32	R. lusitanum p1-7(T) (DQ431674)	0.144	0.098	0.135	0.137	0.144	0.135	0.137	0.140	0.112	0.137	0.171	0.121	0.137	0.137	0.137	0.151	0.171	0.144	0.157	0.171	0.149	0.118	0.152	0.149	0.094	0.145	0.091	0.107	0.162	0.147	0.111														
33	R. mayense CCGE526(T) (JX855195)	0.176	0.082	0.142	0.170	0.176	0.147	0.170	0.171	0.071	0.170	0.198	0.075	0.170	0.170	0.170	0.150	0.177	0.176	0.172	0.180	0.142	0.086	0.161	0.182	0.117	0.148	0.109	0.065	0.147	0.151	0.082	0.122													
34	R. mongolense ATCC BAA-116(T) (AY907358)	0.138	0.109	0.139	0.114	0.138	0.127	0.114	0.134	0.101	0.114	0.151	0.118	0.114	0.114	0.114	0.132	0.126	0.138	0.121	0.141	0.031	0.101	0.139	0.139	0.130	0.028	0.125	0.126	0.143	0.115	0.093	0.128	0.126												
35	R. multihospitium CCBAU 83401(T) (EF490029)	0.135	0.072	0.122	0.110	0.135	0.114	0.110	0.131	0.072	0.110	0.130	0.087	0.110	0.110	0.110	0.155	0.129	0.135	0.125	0.135	0.130	0.093	0.122	0.125	0.079	0.135	0.061	0.111	0.130	0.125	0.076	0.124	0.110	0.118											
36	R. paranaense PRF 35(T) (EU488826)	0.123	0.003	0.102	0.115	0.123	0.111	0.115	0.128	0.037	0.115	0.151	0.028	0.115	0.115	0.115	0.125	0.122	0.123	0.139	0.141	0.116	0.078	0.123	0.122	0.083	0.125	0.068	0.068	0.139	0.122	0.041	0.094	0.078	0.109	0.068										
37	R. phaseoli ATCC 14482(T) (EF113136)	0.000	0.128	0.103	0.091	0.000	0.108	0.091	0.003	0.142	0.091	0.127	0.147	0.091	0.091	0.091	0.149	0.047	0.000	0.102	0.099	0.146	0.143	0.116	0.106	0.126	0.147	0.113	0.168	0.140	0.107	0.138	0.145	0.176	0.139	0.136	0.124									
38	R. pisi DSM 30132(T) (DQ431676)	0.115	0.132	0.116	0.066	0.115	0.112	0.066	0.111	0.119	0.066	0.102	0.133	0.066	0.066	0.066	0.183	0.103	0.115	0.080	0.077	0.163	0.126	0.116	0.027	0.139	0.163	0.126	0.143	0.064	0.058	0.124	0.151	0.170	0.145	0.118	0.128	0.115								
39	R. pusense NRCPB10(T) (HQ166059)	0.136	0.125	0.142	0.137	0.136	0.142	0.137	0.132	0.125	0.137	0.162	0.143	0.137	0.137	0.137	0.054	0.140	0.136	0.140	0.149	0.161	0.146	0.154	0.158	0.117	0.171	0.116	0.169	0.178	0.168	0.121	0.169	0.158	0.144	0.128	0.121	0.136	0.159							
40	R. radiobacter ATCC 19358(T) (AM182121)	0.183	0.152	0.183	0.151	0.183	0.184	0.151	0.178	0.167	0.151	0.200	0.172	0.151	0.151	0.151	0.069	0.198	0.183	0.167	0.187	0.187	0.176	0.198	0.202	0.140	0.198	0.147	0.208	0.227	0.199	0.153	0.180	0.189	0.169	0.161	0.148	0.184	0.215	0.083						
41	R. sophorae CCBAU 03386(T) (KJ831252)	0.109	0.143	0.114	0.068	0.109	0.115	0.068	0.105	0.135	0.068	0.101	0.140	0.068	0.068	0.068	0.159	0.082	0.109	0.071	0.073	0.148	0.125	0.116	0.095	0.133	0.159	0.128	0.142	0.047	0.028	0.144	0.160	0.164	0.141	0.133	0.139	0.110	0.068	0.162	0.202					
42	R. sophoriradicis CCBAU 03470(T) (KJ831248)	0.047	0.126	0.107	0.080	0.047	0.103	0.080	0.044	0.130	0.080	0.106	0.135	0.080	0.080	0.080	0.157	0.000	0.047	0.082	0.084	0.146	0.150	0.103	0.102	0.133	0.152	0.120	0.164	0.110	0.087	0.126	0.171	0.177	0.126	0.129	0.122	0.047	0.103	0.140	0.198	0.082				
43	R. tropici CIAT 899(T) (CP004015)	0.168	0.093	0.158	0.141	0.168	0.159	0.141	0.163	0.097	0.141	0.167	0.097	0.141	0.141	0.141	0.154	0.157	0.168	0.150	0.158	0.162	0.127	0.167	0.155	0.068	0.177	0.050	0.113	0.153	0.149	0.093	0.134	0.120	0.158	0.078	0.089	0.169	0.146	0.141	0.169	0.148	0.157			
44	R. vallis CCBA 65647(T) (GU211770)	0.094	0.109	0.110	0.061	0.094	0.098	0.061	0.090	0.122	0.061	0.090	0.123	0.061	0.061	0.061	0.150	0.082	0.094	0.089	0.099	0.151	0.152	0.101	0.106	0.129	0.157	0.109	0.138	0.119	0.102	0.126	0.123	0.171	0.130	0.122	0.109	0.094	0.103	0.149	0.169	0.114	0.082	0.144		
45	B. japonicum DSMZ30131(T) (AY591555)	0.269	0.270	0.255	0.274	0.269	0.266	0.274	0.263	0.291	0.274	0.300	0.292	0.274	0.274	0.274	0.269	0.251	0.269	0.284	0.315	0.277	0.294	0.271	0.322	0.305	0.296	0.291	0.303	0.342	0.290	0.292	0.297	0.313	0.272	0.297	0.275	0.270	0.330	0.280	0.258	0.296	0.251	0.320	0.278	

Appendix 3: Estimates of evolutionary divergence between *nodC* sequences of study strains and of closely related described species. The number of base substitutions per site from between sequences is shown. Analyses were conducted using the Tamura 3-parameter model. The rate variation among sites was modelled with a gamma distribution (shape parameter = 1). The analysis involved 35 nucleotide sequences. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. There were a total of 489 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013).

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
1	NAK 103																																			
2	NAK 120	0.383																																		
3	NAK 210	0.035	0.418																																	
4	NAK 239	0.000	0.383	0.035																																
5	NAK 245	0.039	0.407	0.025	0.039																															
6	NAK 266	0.039	0.407	0.025	0.039	0.000																														
7	NAK 287	0.000	0.383	0.035	0.000	0.039	0.039																													
8	NAK 294	0.363	0.037	0.397	0.363	0.386	0.386	0.363																												
9	NAK 312	0.039	0.407	0.025	0.039	0.000	0.000	0.039	0.386																											
10	NAK 334	0.039	0.407	0.025	0.039	0.000	0.000	0.039	0.386	0.000																										
11	NAK 349	0.363	0.037	0.397	0.363	0.386	0.386	0.363	0.000	0.386	0.386																									
12	NAK 358	0.039	0.407	0.025	0.039	0.000	0.000	0.039	0.386	0.000	0.000	0.386																								
13	NAK 382	0.035	0.418	0.000	0.035	0.025	0.025	0.035	0.397	0.025	0.025	0.397	0.025																							
14	NAK 387	0.035	0.418	0.000	0.035	0.025	0.025	0.035	0.397	0.025	0.025	0.397	0.025	0.000																						
15	NAK 458	0.000	0.383	0.035	0.000	0.039	0.039	0.000	0.363	0.039	0.039	0.363	0.039	0.035	0.035																					
16	R. acidisoli FH23(T) (KJ921061)	0.026	0.402	0.012	0.026	0.012	0.012	0.026	0.381	0.012	0.012	0.381	0.012	0.012	0.012	0.026																				
17	R. anhuiense CCBAU 23252(T) (KF111957)	0.452	0.491	0.469	0.452	0.499	0.499	0.452	0.458	0.499	0.499	0.458	0.499	0.469	0.469	0.452	0.476																			
18	R. etli CFN 42(T) (U80928)	0.000	0.383	0.035	0.000	0.039	0.039	0.000	0.363	0.039	0.039	0.363	0.039	0.035	0.035	0.000	0.026	0.452																		
19	R. fabae CCBAU 33202(T) (JN580683)	0.463	0.503	0.480	0.463	0.511	0.511	0.463	0.469	0.511	0.511	0.469	0.511	0.480	0.480	0.463	0.488	0.006	0.463																	
20	R. gallicum R602sp(T) (AF217266)	0.216	0.292	0.223	0.216	0.229	0.229	0.216	0.286	0.229	0.229	0.286	0.229	0.223	0.223	0.216	0.219	0.361	0.216	0.370																
21	R. hainanense CCBAU 57015(T) (DQ010039)	0.477	0.507	0.501	0.477	0.505	0.505	0.477	0.467	0.505	0.505	0.467	0.505	0.501	0.501	0.477	0.495	0.581	0.477	0.595	0.422															
22	R. laguerreae FB206(T) (KC608575)	0.454	0.484	0.484	0.454	0.501	0.501	0.454	0.451	0.501	0.501	0.451	0.501	0.484	0.484	0.454	0.479	0.021	0.454	0.015	0.355	0.595														
23	R. leguminosarum ATCC 14480 (FJ895269)	0.562	0.492	0.597	0.562	0.619	0.619	0.562	0.525	0.619	0.619	0.525	0.619	0.597	0.597	0.562	0.591	0.441	0.562	0.440	0.450	0.668	0.407													
24	R. leguminosarum USDA 2370(T) (FJ596038)	0.444	0.505	0.466	0.444	0.471	0.471	0.444	0.472	0.471	0.471	0.472	0.471	0.466	0.466	0.444	0.461	0.055	0.444	0.058	0.337	0.552	0.051	0.389												
25	R. leguminosarum WSM1325 (CP001623)	0.577	0.479	0.615	0.577	0.636	0.636	0.577	0.511	0.636	0.636	0.511	0.636	0.615	0.615	0.577	0.608	0.462	0.577	0.460	0.456	0.675	0.432	0.041	0.408											
26	R. leucaenae HBR12 (JN580662)	0.363	0.037	0.397	0.363	0.386	0.386	0.363	0.000	0.386	0.386	0.000	0.386	0.397	0.397	0.363	0.381	0.458	0.363	0.469	0.286	0.467	0.451	0.525	0.472	0.511										
27	R. lusitanum P1-7(T) (HM852098)	0.371	0.041	0.406	0.371	0.394	0.394	0.371	0.004	0.394	0.394	0.004	0.394	0.406	0.406	0.371	0.390	0.468	0.371	0.479	0.293	0.477	0.461	0.536	0.482	0.522	0.004									
28	R. multihospitium CCBAU 83401(T) (EF050781)	0.477	0.495	0.501	0.477	0.505	0.505	0.477	0.462	0.505	0.505	0.462	0.505	0.501	0.501	0.477	0.495	0.098	0.477	0.101	0.368	0.603	0.102	0.435	0.063	0.443	0.462	0.472								
29	R. phaseoli ATCC 14482(T) (HM441255)	0.000	0.383	0.035	0.000	0.039	0.039	0.000	0.363	0.039	0.039	0.363	0.039	0.035	0.035	0.000	0.026	0.452	0.000	0.463	0.216	0.477	0.454	0.562	0.444	0.577	0.363	0.371	0.477							
30	R. pisi DSM 30132(T) (JQ795195)	0.443	0.494	0.465	0.443	0.469	0.469	0.443	0.471	0.469	0.469	0.471	0.469	0.465	0.465	0.443	0.460	0.060	0.443	0.062	0.336	0.551	0.055	0.379	0.004	0.398	0.471	0.481	0.063	0.443						
31	R. sophorae CCBAU 03386(T) (KJ831243)	0.030	0.402	0.017	0.030	0.017	0.017	0.030	0.381	0.017	0.017	0.381	0.017	0.017	0.017	0.030	0.004	0.477	0.030	0.488	0.219	0.496	0.479	0.600	0.461	0.617	0.381	0.390	0.495	0.030	0.460					
32	R. sophoriradicis CCBAU 03470(T) (KJ831245)	0.035	0.418	0.000	0.035	0.025	0.025	0.035	0.397	0.025	0.025	0.397	0.025	0.000	0.000	0.035	0.012	0.469	0.035	0.480	0.223	0.501	0.484	0.597	0.466	0.615	0.397	0.406	0.501	0.035	0.465	0.017				
33	R. tropici CIAT 899(T) (JN580681)	0.363	0.037	0.397	0.363	0.386	0.386	0.363	0.000	0.386	0.386	0.000	0.386	0.397	0.397	0.363	0.381	0.458	0.363	0.469	0.286	0.467	0.451	0.525	0.472	0.511	0.000	0.004	0.462	0.363	0.471	0.381	0.397			
34	R. vallis CCBAU 65647(T) (GU211769)	0.039	0.407	0.025	0.039	0.000	0.000	0.039	0.386	0.000	0.000	0.386	0.000	0.025	0.025	0.039	0.012	0.499	0.039	0.511	0.229	0.505	0.501	0.619	0.471	0.636	0.386	0.394	0.505	0.039	0.469	0.017	0.025	0.386		
35	B. japonicum USDA 6(T) (AP012206)	0.495	0.540	0.490	0.495	0.521	0.521	0.495	0.525	0.521	0.521	0.525	0.521	0.490	0.490	0.495	0.491	0.581	0.495	0.595	0.468	0.568	0.607	0.567	0.612	0.574	0.525	0.536	0.583	0.495	0.611	0.505	0.490	0.525	0.521	

Appendix 4: A 4-point scale for scoring nodulation in *Phaseolus vulgaris*

	Score
1. Distribution	
*Crown + full laterals	1
Crown + partial laterals	0.75
Full laterals	0.5
Partial laterals	0.25
2. Nodule Size**	
Mainly large (>75%)	1
Large + Medium	0.75
Medium + Small	0.5
Small	0.25
3. Nodule Number	
>201	1
101-200	0.75
21-100	0.5
1-20	0.25
4. Nodule color***	
Pink	1
Pink + White/Green	0.75
Pink+ Green + White	0.5
Green/white	0.25

* Crown-5cm of tap root and major laterals
** Sizes: Large ≥4mm, medium 2-4mm, small ≤2
***Indicative of the main colour

Appendix 5: Rhizobia population changes in soil inoculated separately with NAK 120C, NAK 210C, NAK 287C and CIAT 899C over the experimental period

Target density at inoculation	Nitrogen	Strain	Population at inoculation (rhizobia/g of soil)	Population at planting (rhizobia/g of soil)	Population size at 10 DAP (rhizobia/g of soil)	Population size at 30 DAP (rhizobia/g of soil)
Low (10^2)	0	NAK 120C	nd	$6.65\times10^4\pm1.0\times10^4$ a‡	$1.61\times10^5\pm1.9\times10^3$ ab	$0.91\times10^6\pm1.4\times10^4$ d
	0	NAK 210C	nd	$5.02\times10^4\pm9.3\times10^3$ a	$1.05\times10^5\pm1.0\times10^4$ a	$1.94\times10^5\pm5.0\times10^4$ b
	0	NAK 287C	nd	$4.40\times10^4\pm4.0\times10^3$ a	$4.39\times10^5\pm8.8\times10^4$ b	$4.79\times10^4\pm8.4\times10^4$ a
	0	CIAT 899C	nd	$8.68\times10^4\pm2.2\times10^3$ a	$3.78\times10^5\pm1.1\times10^5$ b	$2.89\times10^6\pm8.1\times10^4$ c
		Mean	**nd**	**$6.19\times10^4\pm6.8\times10^3$**	**$2.71\times10^5\pm6.0\times10^4$**	**$1.01\times10^6\pm4.3\times10^5$**
	1	NAK 120C	nd	$3.07\times10^4\pm6.7\times10^3$	$8.75\times10^4\pm2.0\times10^4$	$3.85\times10^5\pm9.8\times10^4$
	1	NAK 210C	nd	$2.22\times10^4\pm1.1\times10^3$	$5.78\times10^4\pm1.5\times10^3$	$4.86\times10^4\pm7.9\times10^3$ *
High (10^5)	0	NAK 120C	$1.07\times10^5\pm2.8\times10^3$ ab	$0.91\times10^6\pm8.3\times10^4$ a	$1.15\times10^6\pm3.7\times10^4$ a	$1.21\times10^5\pm7.9\times10^4$ a
	0	NAK 210C	$5.60\times10^4\pm1.3\times10^4$ a	$0.83\times10^6\pm2.7\times10^5$ a	$5.35\times10^6\pm4.6\times10^5$ c	$2.75\times10^5\pm1.9\times10^4$ b
	0	NAK 287C	$1.16\times10^5\pm2.8\times10^3$ ab	$2.21\times10^6\pm1.4\times10^5$ ab	$6.28\times10^6\pm8.3\times10^4$ c	$1.57\times10^6\pm4.9\times10^4$ c
	0	CIAT 899C	$2.10\times10^5\pm3.6\times10^4$ b	$3.36\times10^6\pm5.1\times10^5$ b	$3.66\times10^6\pm1.5\times10^4$ b	$5.44\times10^6\pm5.1\times10^4$ d
		Mean	**$1.20\times10^5\pm2.1\times10^4$**	**$1.83\times10^6\pm4.0\times10^5$**	**$4.11\times10^6\pm7.4\times10^5$**	**$2.74\times10^6\pm6.3\times10^5$**
	1	NAK 120C	$1.08\times10^5\pm1.2\times10^3$	$1.13\times10^6\pm1.2\times10^5$	$1.81\times10^6\pm1.7\times10^5$	$5.96\times10^6\pm5.8\times10^5$*
	1	NAK 210C	$5.73\times10^4\pm6.9\times10^3$	$7.18\times10^5\pm1.5\times10^5$	$1.34\times10^6\pm1.2\times10^5$*	$7.21\times10^5\pm1.2\times10^5$*

nd: not determined

‡populations, within a sampling time and within a low or high inoculation, marked with the same letter(s) are not significantly different from each other (HSD, p=0.05)

*Populations are significantly different (ANOVA, P=0.05) from comparable treatment lacking supplementary N

References

Abaidoo, R. C., George, T., Bohlool, B. B., & Singleton, P. W. (1990). Influence of elevation and applied nitrogen on rhizosphere colonization and competition for nodule occupancy by different rhizobial strains on field-grown soybean and common bean. *Canadian Journal of Microbiology, 36*(2), 92-96.

Aguilar, O. M., Riva, O., & Peltzer, E. (2004). Analysis of *Rhizobium etli* and of its symbiosis with wild *Phaseolus vulgaris* supports coevolution in centers of host diversification. *Proceedings of the National Academy of Sciences 101*(37), 13548-13553.

Alberton, O., Kaschuk, G., & Hungria, M. (2006). Sampling effects on the assessment of genetic diversity of rhizobia associated with soybean and common bean. *Soil Biology and Biochemistry, 38*(6), 1298-1307.

Andrade, D. S., Murphy, P. J., & Giller, K. E. (1999). Genetic diversity of rhizobial populations nodulating *Phaseolus vulgaris* in acid soils from Brazil. In E. MartÍnez & G. Hernández (Eds.), *Highlights of Nitrogen Fixation Research* (pp. 291-294). Boston, MA: Springer US.

Andrews, M., Raven, J. A., Sprent, J. I., & Lea, P. J. (2007). Is shoot growth correlated to leaf protein concentration? *Trends in Plant Science, 12*(12), 531-532.

Andrews, M., Sprent, J. I., Raven, J. A., & Eady, D. E. (1999). Relationships between shoot to root ratio, growth and leaf soluble protein concentration of *Pisum sativum*, *Phaseolus vulgaris* and *Triticum aestivum* under different nutrient deficiencies. *Plant, Cell and Environment, 22*(8), 949-958.

Aneja, P., Dai, M., Lacorre, D. A., Pillon, B., & Charles, T. C. (2004). Heterologous complementation of the exopolysaccharide synthesis and carbon utilization phenotypes of *Sinorhizobium meliloti* Rm1021 polyhydroxyalkanoate synthesis mutants. *FEMS Microbiology Letters, 239*(2), 277-283.

Aneja, P., Zachertowska, A., & Charles, T. C. (2005). Comparison of the symbiotic and competition phenotypes of *Sinorhizobium meliloti* PHB synthesis and degradation pathway mutants. *Canadian Journal of Microbiology, 51*(7), 599-604.

Anyango, B., Wilson, K., & Giller, K. (1998). Competition in Kenyan soils between *Rhizobium leguminosarum* biovar *phaseoli* strain Kim5 and *R. tropici* strain CIAT 899 using the *gusA* marker gene. *Plant and Soil, 204*, 69-78.

Anyango, B., Wilson, K. J., Beynon, J. L., & Giller, K. E. (1995). Diversity of rhizobia nodulating *Phaseolus vulgaris* L. in two Kenyan soils with contrasting pHs. *Applied and Environmental Microbiology, 61*(11), 4016-4021.

Ardley, J. K., Reeve, W. G., O'Hara, G. W., Yates, R. J., Dilworth, M. J., & Howieson, J. G. (2013). Nodule morphology, symbiotic specificity and association with unusual rhizobia are distinguishing features of the genus *Listia* within the southern African crotalarioid clade *Lotononis* s.l. *Annals of Botany.*

Ardourel, M., Demont, N., Debellé, F., Maillet, F., de Billy, F., Promé, J. C., Dénarié, J., & Truchet, G. (1994). *Rhizobium meliloti* lipooligosaccharide nodulation factors: different structural requirements for bacterial entry into target root hair cells and induction of plant symbiotic developmental responses. *The Plant Cell, 6*(10), 1357-1374.

Armager, N., Macheret, V., & Laguerre, G. (1997). *Rhizobium gallicum* sp. nov. and *Rhizobium giardinii* sp. nov., from *Phaseolus vulgaris* nodules *International Journal of Systematic and Evolutionary Microbiology, 47*, 996-1006.

Aserse, A. A., Räsänen, L. A., Assefa, F., Hailemariam, A., & Lindström, K. (2012). Phylogeny and genetic diversity of native rhizobia nodulating common bean

(Phaseolus vulgaris L.) in Ethiopia. *Systematic and Applied Microbiology, 35*(2), 120-131.

Asfaw, A., Blair, M. W., & Almekinders, C. (2009). Genetic diversity and population structure of common bean (*Phaseolus vulgaris* L.) landraces from the East African highlands. *Theoretical and Applied Genetics, 120*(1), 1-12.

Baele, M., Baele, P., Vaneechoutte, M., Storms, V., Butaye, P., Devriese, L. A., Verschraegen, G., Gillis, M., & Haesebrouck, F. (2000). Application of tRNA intergenic spacer PCR for identification of *Enterococcus* Species. *Journal of Clinical Microbiology, 38*(11), 4201-4207.

Bala, A., Karanja, N. K., Mazvita, M., Lwimbi, L., Abaidoo, R., & Giller, K. (2011). Production and use of rhizobial inoculants in Africa: N2Africa Project Report.

Bala, A., Murphy, P. J., & Giller, K. E. (2001). Genetic diversity of rhizobia from natural populations varies with the soil dilution sampled. *Soil Biology and Biochemistry, 33*(6), 841-843.

Baligar, V. C., Fageria, N. K., & He, Z. L. (2001). Nutrient use efficiency in plants. *Communications in Soil Science and Plant Analysis, 32*(7-8), 921-950.

Balume, I. K., Keya, O., Karanja, N. K., & Woomer, P. (2015). Shelf-life of legume inoculants in different carrier materials available in East Africa. *African Crop Science Journal, 23*(4), 379-385.

Banba, M., Siddique, A.-B. M., Kouchi, H., Izui, K., & Hata, S. (2001). *Lotus japonicus* forms early senescent root nodules with *Rhizobium etli*. *Molecular Plant-Microbe Interactions, 14*(2), 173-180.

Bates, S., Cashmore, A. M., & Wilkins, B. M. (1998). IncP plasmids are unusually effective in mediating conjugation of *Escherichia coli* and *Saccharomyces cerevisiae:* involvement of the Tra2 mating system. *Journal of Bacteriology, 180*(24), 6538-6543.

Bell, T. (2010). Experimental tests of the bacterial distance-decay relationship. *International Society for Microbial Ecology, 4*(11), 1357-1365.

Bianco, L., Angelini, J., Fabra, A., & Malpassi, R. (2013). Diversity and symbiotic effectiveness of indigenous rhizobia-nodulating *Adesmia bicolor* in soils of Central Argentina. *Current Microbiology, 66*(2), 174-184.

Bitocchi, E., Nanni, L., Bellucci, E., Rossi, M., Giardini, A., Zeuli, P. S., Logozzo, G., Stougaard, J., McClean, P., Attene, G., & Papa, R. (2012). Mesoamerican origin of the common bean (*Phaseolus vulgaris* L.) is revealed by sequence data. *Proceedings of the National Academy of Sciences, 109*(14), E788-E796.

Blanco, A., Sicardi, M., & Frioni, L. (2010). Competition for nodule occupancy between introduced and native strains of *Rhizobium leguminosarum* biovar *trifolii*. *Biology and Fertility of Soils, 46*(4), 419-425.

Bliss, F. A. (1993). Breeding common bean for improved biological nitrogen fixation. *Plant and Soil, 152*(1), 71-79.

Boody, G., & DeVore, B. (2006). Redesigning agriculture. *BioScience, 56*(10), 839-845.

Boogerd, F. C., & van Rossum, D. (1997). Nodulation of groundnut by *Bradyrhizobium*: a simple infection process by crack entry. *FEMS Microbiology Reviews, 21*(1), 5-27.

Boucher, Y., Douady, C. J., Sharma, A. K., Kamekura, M., & Doolittle, W. F. (2004). Intragenomic heterogeneity and intergenomic recombination among haloarchaeal rRNA genes. *Journal of Bacteriology, 186*(12), 3980-3990.

Brear, E. M., Day, D. A., & Smith, P. M. C. (2013). Iron: an essential micronutrient for the legume-rhizobium symbiosis. *Frontiers in Plant Science, 4*, 359.

Bressers, E. (2014). *Nutrient deficiencies and soil fertility constraints for common bean (Phaseolus vulgaris L.) production in the Usambara Mountains, northern Tanzania.* (M. Sc.), Wagenigen University, Wagenigen, Netherlands.

Brodrick, S. J., Amijee, F., Kipe-Nolt, J. A., & Giller, K. E. (1995). Seed analysis as a means of identifying micronutrient deficiencies of *Phaseolus vulgaris* L. in the tropics. *Tropical Agriculture, 72*, 277-284.

Brom, S., García-de los Santos, A., Cervantes, L., Palacios, R., & Romero, D. (2000). In *Rhizobium etli* symbiotic plasmid transfer, nodulation competitivity and cellular growth require interaction among different replicons. *Plasmid, 44*(1), 34-43.

Broughton, W. J., & Dilworth, M. (1971). Control of leghaemoglobin synthesis in snake beans. *Biochemistry Journal, 125*, 1075-1080.

Broughton, W. J., Hanin, M., Relić, B., Kopciñska, J., Golinowski, W., Şimşek, Ş., Ojanen-Reuhs, T., Reuhs, B., Marie, C., Kobayashi, H., Bordogna, B., Le Quéré, A., Jabbouri, S., Fellay, R., Perret, X., & Deakin, W. J. (2006). Flavonoid-inducible modifications to rhamnan O antigens are necessary for *Rhizobium* sp. strain NGR234-legume symbioses. *Journal of Bacteriology, 188*(10), 3654-3663.

Broughton, W. J., Hernández, G., Blair, M., Beebe, S., Gepts, P., & Vanderleyden, J. (2003). Beans (*Phaseolus* spp.) – model food legumes. *Plant and Soil, 252*(1), 55-128.

Bullard, G. K., Roughley, R. J., & Pulsford, D. J. (2005). The legume inoculant industry and inoculant quality control in Australia: 1953-2003. *Australian Journal of Experimental Agriculture, 45*, 127-140.

Burgess, B. K., & Lowe, D. J. (1996). Mechanism of molybdenum nitrogenase. *Chemical Reviews, 96*(7), 2983-3012.

Buttery, B. R., Park, J. S., & van Berkum, P. (1997). Effects of common bean (*Phaseolus vulgaris* L.) cultivar and rhizobium strain on plant growth, seed yield and nitrogen content. *Canadian Journal of Plant Science*(77), 347-351.

Caballero-Mellado, J., & Martinez-Romero, E. (1999). Soil fertilization limits the genetic diversity of *Rhizobium* in nean nodules. *Symbiosis, 26*, 111-121.

Cai, G.-q., Driscoll, B. T., & Charles, T. C. (2000). Requirement for the enzymes acetoacetyl coenzyme A synthetase and poly-3-hydroxybutyrate (PHB) synthase for growth of *Sinorhizobium meliloti* on PHB cycle intermediates. *Journal of Bacteriology, 182*(8), 2113-2118.

Cardoso, J. D., Hungria, M., & Andrade, D. S. (2012). Polyphasic approach for the characterization of rhizobial symbionts effective in fixing N_2 with common bean (*Phaseolus vulgaris* L.). *Applied Microbiology and Biotechnology, 93*(5), 2035-2049.

Castro, S., Carrera, I., & Martínez-Drets, G. (2000). Methods to evaluate nodulation competitiveness between *Sinorhizobium meliloti* strains using melanin production as a marker. *Journal of Microbiological Methods, 41*(2), 173-177.

Catroux, G., Hartmann, A., & Revellin, C. (2001a). Trends in rhizobial inoculant production and use. *Plant and Soil*(230), 21-30.

Catroux, G., Hartmann, A., & Revellin, C. (2001b). Trends in rhizobial inoculant production and use. *Plant and Soil, 230*(1), 21-30.

Cebolla, A., Ruiz-Berraquero, F., & Palomares, A. J. (1993). Stable tagging of *Rhizobium meliloti* with the firefly luciferase gene for environmental monitoring. *Applied and Environmental Microbiology, 59*(8), 2511-2519.

Cermola, M., Fedorova, E., Taté, R., Riccio, A., Favre, R., & Patriarca, E. J. (2000). Nodule invasion and symbiosome differentiation during *Rhizobium etli-Phaseolus vulgaris* symbiosis. *Molecular Plant-Microbe Interactions, 13*(7), 733-741.

Cevallos, M. A., Encarnación, S., Leija, A., Mora, Y., & Mora, J. (1996). Genetic and physiological characterization of a *Rhizobium etli* mutant strain unable to synthesize poly-beta-hydroxybutyrate. *Journal of Bacteriology, 178*(6), 1646-1654.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W., & Prasher, D. (1994). Green flourescent protein as a marker for gene expression. *Science, 263*, 802-804.

Chun, K. T., Edenberg, H. J., Kelley, M. R., & Goebl, M. G. (1997). Rapid amplification of uncharacterized transposon-tagged DNA sequences from genomic DNA. *Yeast, 13*, 233-240.

Colebatch, G., Desbrosses, G., Ott, T., Krusell, L., Montanari, O., Kloska, S., Kopka, J., & Udvardi, M. K. (2004). Global changes in transcription orchestrate metabolic differentiation during symbiotic nitrogen fixation in *Lotus japonicus*. *The Plant Journal, 39*(4), 487-512.

Collins, M. T., Thies, J., & Abbott, L. K. (2002). Diversity and symbiotic effectiveness of *Rhizobium leguminosarum* bv. *trifolii* from pasture soils in south-western Australia. *Australian Journal of Soil Research, 40*, 1319-1329.

Cooper, J. E. (2007). Early interactions between legumes and rhizobia: disclosing complexity in a molecular dialogue. *Journal of Applied Microbiology, 103*(5), 1355-1365.

Corbin, E. J., Brockwell, J., & Gault, R. R. (1977). Nodulation studies on chickpea (*Cicer arietum*). *Australian Journal Experimental Agriculture and Animal Husbandry, 17*, 126-134.

Corby, H. D. L. (1988). The types of rhizobial nodule and their distribution among the Leguminosae. *Kirkia, 13*, 53-123.

Corich, V., Giacomini, A., Vian, P., Vendramin, E., Carlot, M., Basaglia, M., Squartini, A., Casella, S., & Nuti, M. (2001). Aspects of marker/reporter stability and selectivity in soil microbiology. *Microbial Ecology, 41*, 333-340.

Crews, T. E., & Peoples, M. B. (2004). Legume versus fertilizer sources of nitrogen: ecological tradeoffs and human needs. *Agriculture, Ecosystems and Environment, 102*(3), 279-297.

Da Silva, P. M., Tsai, S. M., & Bonetti, R. (1993). Response to inoculation and N fertilization for increased yield and biological nitrogen fixation of common bean (*Phaseolus vulgaris* L.). *Plant and Soil, 152*(1), 123-130.

Dall'Agnol, R. F., Plotegher, F., Souza, R. C., Mendes, I. C., dos Reis Junior, F. B., Béna, G., Moulin, L., & Hungria, M. (2016). *Paraburkholderia nodosa* is the main N_2-fixing species trapped by promiscuous common bean (*Phaseolus vulgaris* L.) in the Brazilian 'Cerradão'. *FEMS Microbiology Ecology, 92*(8).

Dall'Agnol, R. F., Ribeiro, R. A., Ormeno-Orrillo, E., Rogel, M. A., Delamuta, J. R., Andrade, D. S., Martinez-Romero, E., & Hungria, M. (2013). *Rhizobium freirei*, a symbiont of *Phaseolus vulgaris* very effective in fixing nitrogen. *International Journal of Systematic and Evolutionary Microbiology*, 4167-4173.

Dall'Agnol, R. F., Ribeiro, R. A., Delamuta, J. R. M., Ormeño-Orrillo, E., Rogel, M. A., Andrade, D. S., Martínez-Romero, E., & Hungria, M. (2014). *Rhizobium paranaense* sp. nov., an effective N_2-fixing symbiont of common bean (*Phaseolus vulgaris L.*) with broad geographical distribution in Brazil. *International Journal of Systematic and Evolutionary Microbiology, 64*(9), 3222-3229.

Dall'Agnol, R. F., Ribeiro, R. A., Ormeño-Orrill, E., Rogel, M. A., Delamuta, J. R. M., Andrade, D. S., Martínez-Romero, E., & Hungria, M. (2013). *Rhizobium freirei*, a symbiont of *Phaseolus vulgaris* very effective in fixing nitrogen. *International Journal of Systematic and Evolutionary Microbiology, 63*.

De Bruijn, F. J. (1992). Use of repetitive (repetitive extragenic palindromic and enterobacterial repetitive intergeneric consensus) sequences and the polymerase chain reaction to fingerprint the genomes of *Rhizobium meliloti* isolates and other soil bacteria. *Applied and Envronmental Microbiology, 58*, 2180-2187.

De Gelder, L., Ponciano, J. M., Joyce, P., & Top, E. M. (2007). Stability of a promiscuous plasmid in different hosts: no guarantee for a long-term relationship. *Microbiology, 153*(Pt 2), 452-463.

De Lorenzo, V., Herrero, M., Sachez, J. M., & Timmis, K. N. (1998). Mini-transposons in microbial ecology and environmental biotechnology. *FEMS Microbiology Ecology, 27*, 211-224.

Deakin, W. J., & Broughton, W. J. (2009). Symbiotic Use of Pathogenic strategies: rhizobial protein secretion systems. *Nature Reviews, 7*, 312-320.

Degefu, T., Wolde-meskel, E., & Frostegård, Å. (2013). Phylogenetic diversity of *Rhizobium* strains nodulating diverse legume species growing in Ethiopia. *Systematic and Applied Microbiology, 36*(4), 272-280.

Del Cerro, P., Rolla-Santos, A. A. P., Gomes, D. F., Marks, B. B., Pérez-Montaño, F., Rodríguez-Carvajal, M. Á., Nakatani, A. S., Gil-Serrano, A., Megías, M., Ollero, F. J., & Hungria, M. (2015). Regulatory *nodD1* and *nodD2* genes of *Rhizobium tropici* strain CIAT 899 and their roles in the early stages of molecular signaling and host-legume nodulation. *BMC Genomics, 16*(1), 1-13.

Denton, M., Reeve, W. G., Howieson, J. G., & Coventry, D. R. (2003). Competitive abilities of common field isolates and a commercial strain of *Rhizobium leguminosarum* bv. *trifolii* for clover nodule occupancy. *Soil Biology and Biochemistry, 35*(8), 1039-1048.

Denton, M. D., Coventry, D. R., Murphy, P. J., Howieson, J. G., & Bellotti, W. D. (2002). Competition between inoculant and naturalised *Rhizobium leguminosarum* bv. *trifolii* for nodulation of annual clovers in alkaline soils. *Australian Journal of Agricultural Research, 53*(9), 1019-1026.

Dixon, R., & Kahn, D. (2004). Genetic regulation of biological nitrogen fixation. *Nature Reviews Microbiology, 2*, 621-631.

Downie, J. A. (2014). Legume nodulation. *Current Biology, 24*(5), R184-R190.

Duodu, S., Brophy, C., Connolly, J., & Svenning, M. M. (2008). Competitiveness of a native *Rhizobium leguminosarum* biovar *trifolii* strain for nodule occupancy is manifested during infection. *Plant and Soil, 318*(1), 117-126.

Easter, C. L., Schwab, H., & Helinski, D. R. (1998). Role of the *parCBA* operon of the broad-host-range plasmid RK2 in stable plasmid maintenance. *Journal of Bacteriology, 180*(22), 6023-6030.

Estévez, J., Soria-Díaz, M. E., De Córdoba, F. F., Morón, B., Manyani, H., Gil, A., Thomas-Oates, J., Van Brussel, A. A. N., Dardanelli, M. S., Sousa, C., & Megías, M. (2009). Different and new Nod factors produced by *Rhizobium tropici* CIAT 899 following Na^{+}stress. *FEMS Microbiology Letters, 293*(2), 220-231.

Evans, J. R. (1989). Photosynthesis and nitrogen relationships in leaves of C3 plants. *Oecologia, 78*(1), 9-19.

Faghire, M., Mandri, B., Oufdou, K., Bargaz, A., Ghoulam, C., Ramírez-Bahena, M. H., Velázquez, E., & Peix, A. (2012). Identification at the species and symbiovar levels of strains nodulating *Phaseolus vulgaris* in saline soils of the Marrakech region (Morocco) and analysis of the *otsA* gene putatively involved in osmotolerance. *Systematic and Applied Microbiology, 35*(3), 156-164.

FAOSTAT. Retrieved 17/08/2016, from http://faostat3.fao.org/compare/E

Faruque, O. M., Miwa, H., Yasuda, M., Fujii, Y., Kaneko, T., Sato, S., & Okazaki, S. (2015). Identification of *Bradyrhizobium elkanii* genes involved in incompatibility with soybean plants carrying the Rj4 allele. *Applied and Environmental Microbiology, 81*(19), 6710-6717.

Ferguson, B. J., Indrasumunar, A., Hayashi, S., Lin, M. H., Lin, Y. H., Reid, D. E., & Gresshoff, P. M. (2010). Molecular analysis of legume nodule development and autoregulation. *Journal of Integrative Plant Biology, 52*(1), 61-76.

Fernández-López, M., Goormachtig, S., Gao, M., D'Haeze, W., Van Montagu, M., & Holsters, M. (1998). Ethylene-mediated phenotypic plasticity in root nodule

development on *Sesbania rostrata*. *Proceedings of the National Academy of Sciences, 95*(21), 12724-12728.

Ferreira, L., Sánchez-Juanes, F., García-Fraile, P., Rivas, R., Mateos, P. F., Martínez-Molina, E., González-Buitrago, J. M., & Velázquez, E. (2011). MALDI-TOF mass spectrometry is a fast and reliable platform for identification and ecological studies of species from family *Rhizobiaceae*. *PLoS ONE, 6*(5), e20223.

Field, C., & Mooney, H. A. (1986). The photosynthesis-nitrogen relationship in wild plants. In T. J. Givnish (Ed.), *On the economy of plant and function* (pp. 25-55). Cambridge: Cambridge University Press.

Finke, R. L., Harper, J. E., & Hageman, R. H. (1982). Efficiency of nitrogen assimilation by N(2)-fixing and nitrate-grown soybean plants (*Glycine max* [L.] Merr.). *Plant Physiology, 70*(4), 1178-1184.

Fox, G. E., Wisotzkey, J. D., & Jurtshuk, J. P. (1992). How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. *International Journal of Systematic Bacteriology, 42*(1), 166-170.

Franzel, S., Carsan, S., Lukuyu, B., Sinja, J., & Wambugu, C. (2014). Fodder trees for improving livestock productivity and smallholder livelihoods in Africa. *Current Opinion in Environmental Sustainability, 6*, 98-103.

Frederix, M., Edwards, A., Swiderska, A., Stanger, A., Karunakaran, R., Williams, A., Abbruscato, P., Sanchez-Contreras, M., Poole, P. S., & Downie, J. A. (2014). Mutation of *pra*R in *Rhizobium leguminosarum* enhances root biofilms, improving nodulation competitiveness by increased expression of attachment proteins. *Molecular Microbiology, 93*(3), 464-478.

Fujikake, H., Yamazaki, A., Ohtake, N., Sueyoshi, K., Matsuhashi, S., Ito, T., Mizuniwa, C., Kume, T., Hashimoto, S., Ishioka, N. S., Watanabe, S., Osa, A., Sekine, T., Uchida, H., Tsuji, A., & Ohyama, T. (2003). Quick and reversible inhibition of soybean root nodule growth by nitrate involves a decrease in sucrose supply to nodules. *Journal of Experimental Botany, 54*(386), 1379-1388.

Gage, D. J. (2002). Analysis of infection thread development using Gfp- and Dsred-expressing *Sinorhizobium meliloti*. *Journal of Bacteriology, 184*(24), 7042-7046.

Gage, D. J., Bobo, T., & Long, S. R. (1996). Use of Green Flourescent Protein to visualize the early events of symbiosis between *Rhizobium meliloti* and alfafa (*Medicago sativa*). *Journal of Bacteriology, 178*(24), 7159-7166.

Galloway, J. N., Dentener, F. J., Capone, D. G., Boyer, E. W., Howarth, R. W., Seitzinger, S. P., Asner, G. P., Cleveland, C. C., Green, P. A., Holland, E. A., Karl, D. M., Michaels, A. F., Porter, J. H., Townsend, A. R., & Vöosmarty, C. J. (2004). Nitrogen cycles: past, present, and future. *Biogeochemistry, 70*(2), 153-226.

Gaunt, M. W., Turner, S. L., Rigottier-Gois, L., Lloyd-Macgilp, S. A., & Young, J. P. W. (2001). Phylogenies of *atpD* and *recA* support the small subunit rRNA-based classification of rhizobia. *International Journal of Systematic and Evolutionary Microbiology 51*, 2037-2048.

George, M. L. C., Robert, F. M., & Bohlool, B. (1992). Nodulation supression by *Rhizobium leguminosarum* bv. *phaseoli* in bean split-root systems. *Symbiosis, 12*, 95-105.

Gerding, M., O'Hara, G. W., Howieson, J. G., & Bräu, L. (2014). Overcoming non-selective nodulation of *Lessertia* by soil-borne rhizobium in the presence of inoculant mesorhizobium. *Plant and Soil, 380*(1), 117-132.

Gevers, D., Cohan, F. M., Lawrence, J. G., Spratt, B. G., Coenye, T., Feil, E. J., Stackebrandt, E., de Peer, Y. V., Vandamme, P., Thompson, F. L., & Swings, J. (2005). Re-evaluating prokaryotic species. *Nature Reviews Microbiology, 3*, 733-739.

Gevers, D., Huys, G., & Swings, J. (2001). Applicability of rep-PCR fingerprinting for identification of *Lactobacillus* species. *FEMS Microbiology Letters, 205*, 31-36.

Giacomini, A., Ollero, F. J., Squartini, A., & Nuti, M. P. (1994). Construction of multipurpose gene cartridges based on a novel synthetic promoter for high-level gene expression in Gram-negative bacteria. *Gene, 144*(1), 17-24.

Gicharu, G. K., Gitonga, N. M., Boga, H., Cheruiyot, R. C., & Maingi, J. M. (2013). Effects of inoculating selected climbing bean cultivars with different rhizobia strains on nitrogen fixation. *Online International Journal of Microbiology Research, 1*(2), 25-31.

Giller, K. (2001). *Nitrogen Fixation in Tropical Cropping Systems*. Wallingford, GB: CAB International.

Giller, K., Amijee, F., Brodrick, S. J., & Edje, O. T. (1998). Environmental constraints to nodulation and nitrogen fixation of *Phaseolus vulgaris* L. in Tanzania. II. Response to N and P fertilisers and inoculation with *Rhizobium*. *African Crop Science Journal, 6*(2), 171-178.

Glyan'ko, A. K., Vasil'eva, G. G., Mitanova, N. B., & Ishchenko, A. A. (2009). The influence of mineral nitrogen on legume-rhizobium symbiosis. *Biology Bulletin, 36*(3), 250-258.

Goeders, N., & Van Melderen, L. (2014). Toxin-antitoxin systems as multilevel interaction systems. *Toxins, 6*(1), 304.

Gomes, D. F., Ormeno-Orrillo, E., & Hungria, M. (2015). Biodiversity, symbiotic efficiency and genomics of *Rhizobium tropici* and related species. In F. Bruijn (Ed.), *Biological nitrogen fixation*. New Jersey: John Wiley & Sons.

Gordon, A. J., Minchin, F. R., James, C. L., & Komina, O. (1999). Sucrose synthase in legume nodules is essential for nitrogen fixation. *Plant Physiology, 120*(3), 867-878.

Graham, P. H., Draeger, K. J., Ferrey, M. L., Conroy, M. J., Hammer, B. E., Martinez, E., Aarons, S. R., & Quinto, C. (1994). Acid pH tolerance in strains of *Rhizobium* and *Bradyrhizobium*, and initial studies on the basis for acid tolerance of *Rhizobium tropici* UMR1899. *Canadian Journal of Microbiology, 40*(3), 198-207.

Graham, P. H., & Ranalli, P. (1997). Common bean (*Phaseolus vulgaris* L.). *Field Crops Research 53*, 131-146.

Graham, P. H., Rosas, J. C., Estevez de Jensen, C., Peralta, E., Tlusty, B., Acosta-Gallegos, J., & Arraes Pereira, P. A. (2003). Addressing edaphic constraints to bean production: the Bean/Cowpea CRSP project in perspective. *Field Crops Research, 82*(2–3), 179-192.

Graham, P. H., Sadowsky, M. J., Keyser, H. H., Barnet, Y. M., Bradley, R. S., Cooper, J. E., De Ley, D. J., Jarvis, B. D. W., Roslycky, E. B., Strijdom, B. W., & Young, J. P. W. (1991). Proposed minimal standards for the description of new genera and species of root- and stem-nodulating bacteria. *International Journal of Systematic and Evolutionary Microbiology, 41*(4), 582-587.

Greenway, P. J. (1944). The oigin of some east African food plants. *The East African Agricultural Journal, 10*(1), 34-39.

Guasch-Vidal, B., Estévez, J., Dardanelli, M. S., Soria-Díaz, M. E., Córdoba, F. F., Balog, C. I., Manyani, H., Gil-Serrano, A., Thomas-Oates, J., Hensbergen, P. J., Deelder, A. M., Megías, M., & Brussel, A. A. (2013). High NaCl concentrations induce the nod genes of *Rhizobium tropici* CIAT 899 in the absence of flavonoid inducers. *Molecular Plant-Microbe Interactions, 26*.

Hardarson, G., & Atkins, C. (2003). Optimising biological N_2 fixation by legumes in farming systems. *Plant and Soil, 252*(1), 41-54.

Hartwig, A. (1998). The regulation of symbiotic N_2 fixation: a conceptual model of N feedback from the ecosystem to the gene expression level. *Perspectives in Plant Ecology, Evolution and Systematics, 1*(1), 92-120.

Heath, K. D., Burke, P. V., & Stinchcombe, J. R. (2012). Coevolutionary genetic variation in the legume-rhizobium transcriptome. *Molecular Ecology, 21*(19), 4735-4747.

Herrera-Cervera, J. A., Caballero-Mellado, J., Laguerre, G., Tichy, H.-V., Requena, N., Amarger, N., Martínez-Romero, E., Olivares, J., & Sanjuan, J. (1999). At least five rhizobial species nodulate *Phaseolus vulgaris* in a Spanish soil. *FEMS Microbiology Ecology, 30*(1), 87-97.

Herridge, D. F., Brock, P., Muir, S., & Simmons, A. (2015). *The C footprints of 2-year cropping sequences in south-eastern Australia vary according to the relative inputs of legume fixed N, fertiliser N and soil organic N.* Paper presented at the 19th International Congress on Nitrogen Fixation, California, USA.

Herridge, D. F., Hartley, E., & Gemell, L. G. (2014). Rhizobial counts in peat inoculants vary amongst legume inoculant groups at manufacture and with storage: implications for quality standards. *Plant and Soil, 380*(1), 327-336.

Herridge, D. F., Peoples, M. B., & Boddey, R. M. (2008). Global inputs of biological nitrogen fixation in agricultural systems. *Plant and Soil, 311*(1-2), 1-18.

Herridge, D. F., & Rose, I. (2000). Breeding for enhanced nitrogen fixation in crop legumes. *Field Crops Research, 65*, 229±248.

Howieson, J. G., & Ballard, R. A. (2004). Optimising the legume symbiosis in stressful and competitive environments within southern Australia-some contemporary thoughts. *Soil Biology and Biochemistry, 36*, 1261-1273.

Howieson, J. G., Loi, A., & Carr, S. J. (1995). *Biserrula pelecinus* L. - a legume pasture species with potential for acid, duplex soils which is nodulated by unique root-nodule bacteria. *Australian Journal of Agricultural Research, 46*, 997-1009.

Howieson, J. G., Yates, R. J., O'Hara, G. W., Ryder, M., & Real, D. (2005). The interactions of *Rhizobium leguminosarum* biovar *trifolii* in nodulation of annual and perennial *Trifolium* spp. from diverse centres of origin. *Australian Journal of Experimental Agriculture, 45*(3), 199-207.

Hume, D. J., & Blair, D. H. (1992). Effect of numbers of *Bradyrhizobium japonicum* applied in commercial inoculants on soybean seed yield in Ontario. *Canadian Journal of Microbiology, 38*(6), 588-593.

Hungria, M., Andrade, D. d. S., Chueire, L. M. d. O., Probanza, A., Guttierrez-Mañero, F. J., & Megías, M. (2000). Isolation and characterization of new efficient and competitive bean (*Phaseolus vulgaris* L.) rhizobia from Brazil. *Soil Biology and Biochemistry, 32*(11–12), 1515-1528.

Hungria, M., Campo, R. J., & Mendes, I. C. (2003). Benefits of inoculation of the common bean (*Phaseolus vulgaris*) crop with efficient and competitive *Rhizobium tropici* strains. *Biology and Fertility of Soils, 39*(2), 88-93.

Hungria, M., Campo, R. J., Mendes, I. C., & Graham, P. H. (2006). Contribution of biological nitrogen fixation to the N nutrition of grain crops in the tropics: the success of soybean (*Glycine max* L. Merr.) in South America. In R. P. Singh, N. Shankar & P. K. Jaiwal (Eds.), *Nitrogen nutrition and sustainable plant productivity* (pp. 43-93). Houston, Texas: Studium Press.

Hungria, M., Johnston, A. W. B., & Phillips, D. A. (1992). Effects of flavonoids released naturally from bean (*Phaseolus vulgaris*) on nodD-regulated gene transcription in *Rhizobium leguminosarum* bv *phaseoli*. *Molecular Plant-Microbe Interactions, 5*(3), 199-203.

Hungria, M., Joseph, C. M., & Phillips, D. A. (1991). Anthocyanidins and flavonols, major nod gene inducers from seeds of a black-seeded common bean (*Phaseolus vulgaris* L.). *Plant Physiol, 97*, 751-758.

Hungria, M., & Kaschuk, G. (2014). Regulation of N_2 fixation and NO_3^-/NH_4^+ assimilation in nodulated and N-fertilized *Phaseolus vulgaris* L. exposed to high temperature stress. *Environmental and Experimental Botany, 98*, 32-39.

Hungria, M., Loureiro, M. F., Mendes, I. C., Campo, R. J., & Graham, P. H. (2005). Inoculant preparation, production and application. In W. E. Newton (Ed.), *Nitrogen fixation: origins, applications an research progress*. Dordrech: Springer.

Hungria, M., O'Hara, G. W., Zilli, J. E., Araujo, R. S., Deaker, R., & Howieson, J. G. (2016). Isolation and growth of rhizobia. In J. G. Howieson & M. Dilworth (Eds.), *Working with rhizobia* (pp. 39-60). Canberra, Australia: Australian Centre for International Agricultural Research.

Hungria, M., & Vargas, M. (2000). Environmental factors affecting N_2 fixation in grain legumes in the tropics, with an emphasis on Brazil. *Field Crops Research, 65*, 151-164.

Jain, A., & Srivastava, P. (2013). Broad host range plasmids. *FEMS Microbiology Letters, 348*(2), 87-96.

Janczarek, M., Rachwał, K., Cieśla, J., Ginalska, G., & Bieganowski, A. (2015a). Production of exopolysaccharide by *Rhizobium leguminosarum* bv. *trifolii* and its role in bacterial attachment and surface properties. *Plant and Soil, 388*(1), 211-227.

Janczarek, M., Rachwal, K., Marzec, A., Grzadziel, J., & Palusińska-Szysz, M. (2015b). Signal molecules and cell-surface components involved in early stages of the legume-rhizobium interactions. *Appl Soil Ecol, 85*, 94-113.

Jefferson, R. A., Burgess, S. M., & Hirsh, D. (1986). beta-Glucuronidase from *Escherichia coli* as a gene-fusion marker. *Proceedings of the National Academy of Sciences, 83*(22), 8447-8451.

Jiang, Y., Pogliano, J., Helinski, D. R., & Konieczny, I. (2002). ParE toxin encoded by the broad-host-range plasmid RK2 is an inhibitor of *Escherichia coli* gyrase. *Molecular Microbiology, 44*(4), 971-979.

Jiao, Y. S., Yan, H., Ji, Z. J., Liu, Y. H., Sui, X. H., Wang, E. T., Guo, B. L., Chen, W. X., & Chen, W. F. (2015). *Rhizobium sophorae* sp. nov. and *Rhizobium sophoriradicis* sp. nov., nitrogen-fixing rhizobial symbionts of the medicinal legume *Sophora flavescens*. *International Journal of Systematic and Evolutionary Microbiology, 65*(2), 497-503.

Johnston, A. W. B., Hombrecher, G., Brewin, N. J., & Cooper, M. C. (1982). Two transmissible plasmids in *Rhizobium leguminosarum* strain 300. *Journal of General Microbiology, 128*, 85-93.

Kamst, E., Pilling, J., Raamsdonk, L. M., Lugtenberg, B. J., & Spaink, H. P. (1997). *Rhizobium* nodulation protein NodC is an important determinant of chitin oligosaccharide chain length in Nod factor biosynthesis. *Journal of Bacteriology, 179*(7), 2103-2108.

Kannenberg, E. L., & Carlson, R. W. (2001). Lipid A and O-chain modifications cause *Rhizobium* lipopolysaccharides to become hydrophobic during bacteroid development. *Molecular Microbiology, 39*(2), 379-392.

Karanja, N. K., & Wood, M. (1988). Selecting *Rhizobium phaseoli* strains for use with beans (*Phaseolus vulgaris* L.) in Kenya: Infectiveness and tolerance of acidity and aluminium. *Plant and Soil, 112*(1), 7-13.

Karlsson, R., Gonzales-Siles, L., Boulund, F., Svensson-Stadler, L., Skovbjerg, S., Karlsson, A., Davidson, M., Hulth, S., Kristiansson, E., & Moore, E. R. B.

(2015). Proteotyping: Proteomic characterization, classification and identification of microorganisms – A prospectus. *Systematic and Applied Microbiology, 38*(4), 246-257.

Karunakaran, R., Mauchline, T. H., Hosie, A. H. F., & Poole, P. S. (2005). A family of promoter probe vectors incorporating autofluorescent and chromogenic reporter proteins for studying gene expression in Gram-negative bacteria. *Microbiology, 151*(10), 3249-3256.

Kaschuk, G., Kuyper, T. W., Leffelaar, P. A., Hungria, M., & Giller, K. E. (2009). Are the rates of photosynthesis stimulated by the carbon sink strength of rhizobial and arbuscular symbioses? *Soil Biology and Biochemistry, 41*, 1233-1244.

Katungi, E., Farrow, A., Chianu, J., Sperling, L., & Beebe, S. (2009). Common bean in Eastern and Southern Africa: a situation and outlook analysis *Targeting breeding and delivery efforts to improve the livelihoods of the poor in drought prone areas*: International Center for Tropical Agriculture.

Kawaka, F., Dida, M. M., Opala, P. A., Ombori, O., Maingi, J., Osoro, N., Muthini, M., Amoding, A., Mukaminega, D., & Muoma, J. (2014). Symbiotic efficiency of native rhizobia nodulating common bean (*Phaseolus vulgaris* L.) in soils of western Kenya. *International Scholarly Research Notices, 2014*, 1-8.

Keino, L., Baijukya, F., Ng'etich, W., Otinga, A. N., Okalebo, J. R., Njoroge, R., & Mukalama, J. (2016). Nutrients limiting soybean (*Glycine max* l) growth in acrisols and ferralsols of western Kenya. *PLoS ONE, 10*(12), e0145202.

Kelly, J. D., Adams, M. W., & Varner, G. V. (1987). Yield stability of determinate and indeterminate dry bean cultivars. *Theoretical and Applied Genetics, 74*(4), 516-521.

Kim, O.-S., Cho, Y.-J., Lee, K., Yoon, S.-H., Kim, M., Na, H., Park, S.-C., Jeon, Y. S., Lee, J.-H., Yi, H., Won, S., & Chun, J. (2012). Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *International Journal of Systematic and Evolutionary Microbiology, 62*(3), 716-721.

Klappenbach, J. A., Saxman, P. R., Cole, J. R., & Schmidt, T. M. (2001). rrndb: the ribosomal RNA operon copy number database. *Nucleic Acids Research, 29*(1), 181-184.

Koinange, E. M. K., & Gepts, P. (1992). Hybrid weakness in wild *Phaseolus vulgaris* L. *Journal of Heredity, 83*(2), 135-139.

Koinange, M. K. (2015). *Influence of biochar amendment on the effectiveness of elite Kenyan rhizobia nodulating common bean (Phaseolus vulgaris L.).* The Univesity of Nairobi, Nairobi.

Kouchi, H., & Yoneyama, T. (1986). Metabolism of [13C]-labelled photosynthate in plant cytosol and bacteroids of root nodules of *Glycine max*. *Physiologia Plantarum, 68*(2), 238-244.

Krishnan, H. B., & Pueppke, S. G. (1992). A *nolC-lacZ* gene fusion in *Rhizobium fredii* facilitates direct assessment of competition for nodulation of soybean. *Canadian Journal of Microbiology, 38*(6), 515-519.

Kumar, N., Lad, G., Giuntini, E., Kaye, M. E., Udomwong, P., Shamsani, N. J., Young, J. P. W., & Bailly, X. (2015). Bacterial genospecies that are not ecologically coherent: population genomics of *Rhizobium leguminosarum*. *Open Biology, 5*(1).

Kvint, K., Nachin, L., Diez, A., & Nystrom, T. (2003). The bacterial universal stress protein: function and regulation. *Currrent Opinion in Microbiology, 6*, 140-145.

Laeremans, T., & Vanderleyden, J. (1998). Review: Infection and nodulation signalling in *Rhizobium-Phaseolus vulgaris* symbiosis. *World Journal of Microbiology & Biotechnology 14*, 787-808.

Laguerre, G., Louvrier, P., Allard, M.-R., & Amarger, N. (2003). Compatibility of rhizobial genotypes within natural populations of *Rhizobium leguminosarum* Biovar *viciae* for nodulation of host legumes. *Applied and Environmental Microbiology, 69*(4), 2276-2283.

Laguerre, G., Mavingui, P., Allard, M., Charnay, M., Louvrier, P., Mazurier, S., Rigottier-Gois, L., & Amarger, N. (1996). Typing of rhizobia by PCR DNA fingerprinting and PCR-restriction fragment length polymorphism analysis of chromosomal and symbiotic gene regions: Application to *Rhizobium leguminosarum* and its different biovars. *Applied and Environmental Microbiology, 62*(6), 2029-2036.

Lancaster, K. M., Roemelt, M., Ettenhuber, P., Hu, Y., Ribbe, M. W., Neese, F., Bergmann, U., & DeBeer, S. (2011). X-ray emission spectroscopy evidences a central carbon in the nitrogenase iron-molybdenum cofactor. *Science, 334*(6058), 974-977.

Lane, D. J. (1991). 16S/23S rRNA sequencing. In E. Stackebrandt & M. Goodfellow (Eds.), *Nucleic Acids Techniques in Bacterial Systematics* (pp. 115-147). Chichester: John Wiley & Sons.

Langwerden, J. E. (2014). *Soil fertility constraints for the production of common bean (Phaseolus vulgaris L.) in the Usambara Mountains of northern Tanzania.* (M.Sc.), Wagenigen University, Wagenigen, Netherlands.

Lapage, S. P., Sneath, P. H. A., Lessel, E. F., Skerman, V. B. D., Seeliger, H. P. R., & Clark, W. A. (1992). *International Code of Nomenclature of Bacteria*. Washington (DC): ASM Press.

Lavin, M., Herendeen, P. S., & Wojciechowski, M. F. (2005). Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the tertiary. *Systematic Biology, 54*(4), 575-594.

Layzell, D. B., Rainbird, R. M., Atkins, C. A., & Pate, J. S. (1979). Economy of photosynthate use in nitrogen-fixing legume nodules. *Plant Physiology, 64*, 888-891.

Leidi, E. O., & Rodriguez-Navarro, D. N. (2000). Nitrogen and phosphorus availability limit N fixation in bean. *New Phytologist, 147*, 337–346.

Leonard, R. L., Judd, F. W., & Statler, R. (2015). The biology of the subtropical and pantropical shrub, *Sophora tomentosa* L., (Fabaceae), in coastal dune environments. . *International Biology Review, 2*(14), 1-16.

Li, D.-M., & Alexander, M. (1986). Bacterial growth rates and competition affect nodulation and root colonization by *Rhizobium meliloti*. *Applied and Environmental Microbiology, 52*(4), 807-811.

Li, W., Raoult, D., & Fournier, P. (2009). Bacterial strain typing in the genomic era. *FEMS Microbiology Reviews, 33*, 892-916.

Limpens, E., Franken, C., Smit, P., Willemse, J., Bisseling, T., & Geurts, R. (2003). LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. *Science, 302*(5645), 630-633.

Lodwig, E. M., Leonard, M., Marroqui, S., Wheeler, T. R., Findlay, K., Downie, J. A., & Poole, P. (2005). Role of polyhydroxybutyrate and glycogen as carbon storage compounds in pea and bean bacteroids. *The American Phytopathological Society, 18*(1), 67-74.

Lodwig, E. M., & Poole, P. (2010). Metabolism in *Rhizobium* bacteroids. *Critical Reviews in Plant Sciences, 22*(1), 37-78.

Long, S. R. (1996). *Rhizobium* symbiosis: Nod factors in perspective *The Plant Cell, 8*, 1885-1898.

López-García, S. L., Vázquez, T. E. E., Favelukes, G., & Lodeiro, A. R. (2002). Rhizobial position as a main determinant in the problem of competition for nodulation in soybean. *Environmental Microbiology, 4*(4), 216-224.

Lopez-Guerrero, M. G., Ormeno-Orrillo, E., Velazquez, E., Rogel, M. A., Acosta, J. L., Gonzalez, V., Martinez, J., & Martinez-Romero, E. (2012). *Rhizobium etli* taxonomy revised with novel genomic data and analyses. *Systematic and Applied Microbiology, 35*(6), 353-358.

López-López, A., Rogel-Hernández, M. A., Barois, I., Ortiz Ceballos, A. I., Martínez, J., Ormeño-Orrillo, E., & Martínez-Romero, E. (2012). *Rhizobium grahamii* sp. nov., from nodules of *Dalea leporina*, *Leucaena leucocephala* and *Clitoria ternatea*, and *Rhizobium mesoamericanum* sp. nov., from nodules of *Phaseolus vulgaris*, siratro, cowpea and *Mimosa pudica*. *International Journal of Systematic and Evolutionary Microbiology, 62*(9), 2264-2271.

López Díez, L. I., Sánchez Juanes, F., Ramírez Bahena, M. H., González Buitrago, J. M., Peix, A., & Velázquez, E. (2016). Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) analysis of rhizobia nodulating *Phaseolus vulgaris* in different soils. In F. González-Andrés & E. James (Eds.), *Biological Nitrogen Fixation and Beneficial Plant-Microbe Interaction* (pp. 73-83). Cham: Springer International Publishing.

Lowther, W. L., & Patrick, H. N. (1993). Spread of *Rhizobium* and *Bradyrhizobium* in soil. *Soil Biology and Biochemistry, 25*(5), 607-612.

Ludwig, W. (2007). Nucleic acid techniques in bacterial systematics and identification. *International Journal of Food Microbiology, 120*, 225-236.

Lupwayi, N. Z., Olsen, P. E., Sande, E. S., Keyser, H. H., Collins, M. M., Singleton, P. W., & Rice, W. A. (2000). Inoculant quality and its evaluation. *Field Crops Research, 65*(2–3), 259-270.

Madsen, E. B., Madsen, L. H., Radutoiu, S., Olbryt, M., Rakwalska, M., Szczyglowski, K., Sato, S., Kaneko, T., Tabata, S., Sandal, N., & Stougaard, J. (2003). A receptor kinase gene of the LysM type is involved in legumeperception of rhizobial signals. *Nature, 425*(6958), 637-640.

Mamidi, S., Rossi, M., Moghaddam, S. M., Annam, D., Lee, R., Papa, R., & McClean, P. E. (2013). Demographic factors shaped diversity in the two gene pools of wild common bean *Phaseolus vulgaris* L. *Heredity, 110*(3), 267-276.

Marinus, W. (2015). *Opportunities and constraints for climbing bean (Phaseolus vulgaris L.) cultivation by smallholder farmers in the Ugandan highlands: developing a basket of options.* Wagenigen University, Wagenigen, Netherlands.

Maróti, G., & Kondorosi, É. (2014). Nitrogen-fixing *Rhizobium*-legume symbiosis: are polyploidy and host peptide-governed symbiont differentiation general principles of endosymbiosis? *Frontiers in Microbiology, 5*, 326.

Martens, M., Delaere, M., Coopman, R., De Vos, P., Gillis, M., & Willems, A. (2007). Multilocus sequence analysis of *Ensifer* and related taxa. *International Journal of Systematic and Evolutionary Microbiology, 57*(3), 489-503.

Martínez-Molina, E., Sánchez Juanes, F., Carro, L., Flores-Félix, J. D., Martínez-Hidalgo, P., Cerda Castillo, E., González Buitrago, J. M., & Velázquez, E. (2016). Identification of rhizobial strains nodulating *Pisum sativum* in northern Spain soils by MALDI-TOF MS (Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry) analysis. In F. González-Andrés & E. James (Eds.), *Biological Nitrogen Fixation and Beneficial Plant-Microbe Interaction* (pp. 37-44). Cham: Springer International Publishing.

Martinez-Romero, E. (2003). Diversity of *Rhizobium-Phaseolus vulgaris* symbiosis: overview and perspectives. *Plant and Soil, 252*, 11-23.

Martinez-Romero, E., Segovia, L., Mercante, F. M., Francho, A. A., Graham, P. H., & Pardo, M. A. (1991). *Rhizobium tropici*, a novel species nodulating *Phaseolus vulgaris* L. beans and *Leucaena sp*. trees. *International Journal of Systematic Bacteriology, 41*(3), 417-426.

Martínez, E., Palacios, R., & Sánchez, F. (1987). Nitrogen-fixing nodules induced by *Agrobacterium tumefaciens* harboring *Rhizobium phaseoli* plasmids. *Journal of Bacteriology, 169*(6), 2828-2834.

Martiny, J. B. H., Eisen, J. A., Penn, K., Allison, S. D., & Horner-Devine, M. C. (2011). Drivers of bacterial β-diversity depend on spatial scale. *Proceedings of the National Academy of Sciences, 108*(19), 7850-7854.

Masclaux-Daubresse, C., Daniel-Vedele, F., Dechorgnat, J., Chardon, F., Gaufichon, L., & Suzuki, A. (2010). Nitrogen uptake, assimilation and remobilization in plants: challenges for sustainable and productive agriculture. *Annals of Botany, 105*(7), 1141-1157.

Mathis, R., Van Gijsegem, F., De Rycke, R., D'Haeze, W., Van Maelsaeke, E., Anthonio, E., Van Montagu, M., Holsters, M., & Vereecke, D. (2005). Lipopolysaccharides as a communication signal for progression of legume endosymbiosis. *Proceedings of the National Academy of Sciences of the United States of America, 102*(7), 2655-2660.

Michiels, J., Dombrecht, B., Vermeiren, N., Xi, C., Luyten, E., & Vanderleyden, J. (1998). *Phaseolus vulgaris* is a non-selective host for nodulation. *FEMS Microbiology Ecology, 26*(3), 193-205.

Minchin, F. R., & Witty, J. F. (2005). Respiratory/carbon costs of symbiotic nitrogen fixation in legumes. In H. Lambers & M. Ribas-Carbo (Eds.), *Plant respiration: from cell to ecosystem* (pp. 195-205). Dordrecht, The Netherlands: Springer.

Mnasri, B., Liu, T. Y., Saidi, S., Chen, W. F., Chen, W. X., Zhang, X. X., & Mhamdi, R. (2014). *Rhizobium azibense* sp. nov., a nitrogen fixing bacterium isolated from root-nodules of *Phaseolus vulgaris*. *International Journal of Systematic and Evolutionary Microbiology, 64*(5), 1501-1506.

Mnasri, B., Mrabet, M., Laguerre, G., Aouani, M. E., & Mhamdi, R. (2007). Salt-tolerant rhizobia isolated from a Tunisian oasis that are highly effective for symbiotic N_2-fixation with *Phaseolus vulgaris* constitute a novel biovar (bv. mediterranense) of *Sinorhizobium meliloti*. *Archives of Microbiology, 187*(1), 79-85.

Mnasri, B., Saïdi, S., Chihaoui, S.-A., & Mhamdi, R. (2012). *Sinorhizobium americanum* symbiovar mediterranense is a predominant symbiont that nodulates and fixes nitrogen with common bean (*Phaseolus vulgaris* L.) in a Northern Tunisian field. *Systematic and Applied Microbiology, 35*(4), 263-269.

Montealegre, C., & Graham, P. H. (1996). Preference in the nodulation of Phaseolus vulgaris cv. RAB39. II. Effect of delayed inoculation or low cell representation in the inoculant on nodule occupancy by Rhizobium tropici UMR1899. *Canadian Journal of Microbiology, 42*(8), 844-850.

Montealegre, C., Graham, P. H., & Kipe-Nolt, J. A. (1995). Preference in the nodulation of *Phaseolus vulgaris* cultivar RAB39. *Canadian Journal of Microbiology, 41*(11), 992-998.

Mora, Y., Díaz, R., Vargas-Lagunas, C., Peralta, H., Guerrero, G., Aguilar, A., Encarnación, S., Girard, L., & Mora, J. (2014). Nitrogen-fixing rhizobial strains isolated from common bean seeds: Phylogeny, physiology, and genome analysis. *Applied and Environmental Microbiology, 80*(18), 5644-5654.

Morón, B., Soria-Díaz, M. E., Ault, J., Verroios, G., Noreen, S., Rodríguez-Navarro, D. N., Gil-Serrano, A., Thomas-Oates, J., Megías, M., & Sousa, C. (2005). Low pH changes the profile of nodulation factors produced by *Rhizobium tropici* CIAT 899. *Chem Biol, 12*.

Mortier, V., Holsters, M., & Goormachtig, S. (2012). Never too many? How legumes control nodule numbers. *Plant, Cell & Environment, 35*(2), 245-258.

Mostasso, L., Mostasso, F. L., Dias, B. G., Vargas, M. A. T., & Hungria, M. (2002). Selection of bean (*Phaseolus vulgaris* L.) rhizobial strains for the Brazilian Cerrados. *Field Crops Research, 73*(2–3), 121-132.

Mrabet, M., Mnasri, B., Romdhane, S. B., Laguerre, G., Aouani, M. E., & Mhamdi, R. (2006). *Agrobacterium* strains isolated from root nodules of common bean specifically reduce nodulation by *Rhizobium gallicum*. *FEMS Microbiology Ecology, 56*(2), 304-309.

Müller, S., Pereira, P. A. A., & Martin, P. (1993). Effect of different levels of mineral nitrogen on nodulation and N_2 fixation of two cultivars of common bean (*Phaseolus vulgaris* L.). *Plant and Soil, 152*(1), 139-143.

Mus, F., Crook, M. B., Garcia, K., Garcia Costas, A., Geddes, B. A., Kouri, E. D., Paramasivan, P., Ryu, M.-H., Oldroyd, G. E. D., Poole, P. S., Udvardi, M. K., Voigt, C. A., Ané, J.-M., & Peters, J. W. (2016). Symbiotic nitrogen fixation and the challenges to its extension to nonlegumes. *Applied and Environmental Microbiology, 82*(13), 3698-3710.

Musandu, A. A., & Ogendo, O. J. (2001). Response of common bean to *Rhizobium* inoculation and fertilisers. *The Journal of Food Technology in Africa, 6*(4), 121-125.

Muthini, M., Maingi, J., Muoma, J., Amoding, A., Mukaminega, D., Osoro, N., Mgutu, A., & Ombori, O. (2014). Morphological assessment and effectiveness of indigenous rhizobia isolates that nodulate *P. vulgaris* in water hyacinth compost testing field in Lake Victoria basin. *British Journal of Applied Science and Technology, 4*(5), 718-738.

Mwenda, G. M., Karanja, N. K., Boga, H., Kahindi, J. H. P., Muigai, A., & Odee, D. W. (2011). Abundance and diversity of legume nodulating rhizobia in soils of Embu district, Kenya. *Tropical and Subtropical Agroecosystems, 13*, 1-10.

Nandasena, K. G., O'Hara, G. W., Tiwari, R. P., Sezmiş, E., & Howieson, J. G. (2007). *In situ* lateral transfer of symbiosis islands results in rapid evolution of diverse competitive strains of mesorhizobia suboptimal in symbiotic nitrogen fixation on the pasture legume *Biserrula pelecinus* L. *Environmental Microbiology, 9*(10), 2496-2511.

Nekola, J. C., & White, P. S. (1999). The distance decay of similarity in biogeography and ecology. *Journal of Biogeography, 26*(4), 867-878.

Niehaus, K., Lagares, A., & Pühler, A. (1998). A *Sinorhizobium meliloti* lipopolysaccharide mutant induces effective nodules on the host plant *Medicago sativa* (alfalfa) but fails to establish a symbiosis with *Medicago truncatula*. *Molecular Plant-Microbe Interactions, 11*(9), 906-914.

Nyström, T., & Neidhardt, F. C. (1993). Isolation and properties of a mutant of *Escherichia coli* with an insertional inactivation of the *uspA* gene, which encodes a universal stress protein. *Journal of Bacteriology, 175*(13), 3949-3956.

O'Hara, G. W. (2001). Nutritional constraints on root nodule bacteria affecting symbiotic nitrogen fixation: a review. *Australian Journal of Experimental Agriculture, 41*(3), 417-433.

O'Hara, G. W., Hungria, M., Woomer, P., & Howieson, J. G. (2016). Counting rhizobia. In J. G. Howieson & M. Dilworth (Eds.), *Working with rhizobia* (pp. 109-124). Canberra, Australia: Australian Centre for International Agricultural Research.

Odee, D. W., Haukka, K., McInroy, S. G., Sprent, J. I., Sutherland, J. M., & Young, J. P. W. (2002). Genetic and symbiotic characterization of rhizobia isolated from tree and herbaceous legumes grown in soils from ecologically diverse sites in Kenya. *Soil Biology and Biochemistry, 34*(6), 801-811.

Oldroyd, G. E., & Downie, J. A. (2008). Coordinating nodule morphogenesis with rhizobial infection in legumes. *Annual review of Plant Biology, 59*, 519-546.

Oldroyd, G. E., Murray, J. D., Poole, P. S., & Downie, J. A. (2011). The rules of engagement in the legume-rhizobial symbiosis. *Annual Review of Genetics, 45*, 119-144.

Oldroyd, G. E. D. (2013). Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nature Reviews Microbiology, 11*(4), 252-263.

Oliveira, J. P., Galli-Terasawa, L. V., Enke, C. G., Cordeiro, V. K., Armstrong, L. C. T., & Hungria, M. (2011). Genetic diversity of rhizobia in a Brazilian oxisol nodulating Mesoamerican and Andean genotypes of common bean (*Phaseolus vulgaris* L.). *World Journal of Microbiology and Biotechnology, 27*(3), 643-650.

Ormeño-Orrillo, E., Menna, P., Almeida, L. G. P., Ollero, F. J., Nicolás, M. F., Rodrigues, E. P., Nakatani, A. S., Batista, J. S. S., Chueire, L. M. O., Souza, R. C., Vasconcelos, A. T. R., Megías, M., Hungria, M., & Martínez-Romero, E. (2012). Genomic basis of broad host range and environmental adaptability of *Rhizobium tropici* CIAT 899 and *Rhizobium sp.* PRF 81 which are used in inoculants for common bean (*Phaseolus vulgaris* L.). *BMC Genomics, 13*(735).

Pacios-Bras, C., van der Burgt, Y. E. M., Deelder, A. M., Vinuesa, P., Werner, D., & Spaink, H. P. (2002). Novel lipochitin oligosaccharide structures produced by *Rhizobium etli* KIM5s. *Carbohydrate Research, 337*(13), 1193-1202.

Panday, D., Schumann, P., & Das, S. K. (2011). *Rhizobium pusense* sp. nov., isolated from the rhizosphere of chickpea (*Cicer arietinum* L.). *International Journal of Systematic and Evolutionary Microbiology, 61*(11), 2632-2639.

Park, S. J., & Buttery, B. R. (1989). Identification and characterization of common bean (*Phaseolus vulgaris* L.) lines well nodulated in the presence of high nitrate. *Plant and Soil, 119*(2), 237-244.

Parker, C. T., Tindall, B. J., & Garrity, G. M. (2015). International Code of Nomenclature of Prokaryotes. *International Journal of Systematic and Evolutionary Microbiology.*

Patrick, H. N., & Lowther, W. L. (1995). Influence of the number of rhizobia on the nodulation and establishment of *Trifolium ambiguum*. *Soil Biology and Biochemistry, 27*(4), 717-720.

Peck, M. C., Fisher, R. F., & Long, S. R. (2006). Diverse flavonoids stimulate NodD1 binding to *nod* gene promoters in *Sinorhizobium meliloti*. *Journal of Bacteriology, 188*(15), 5417-5427.

Pellock, B. J., Cheng, H.-P., & Walker, G. C. (2000). Alfalfa root nodule invasion efficiency is dependent on *Sinorhizobium meliloti* polysaccharides. *Journal of Bacteriology, 182*(15), 4310-4318.

Peoples, Brockwell, J., Herridge, D. F., Rochester, I. J., Alves, B. J. R., Urquiaga, S., Boddey, R. M., Dakora, F. D., Bhattarai, S., Maskey, S. L., Sampet, C., Rerkasem, B., Khan, D. F., Hauggaard-Nielsen, H., & Jensen, E. S. (2009). The contributions of nitrogen-fixing crop legumes to the productivity of agricultural systems. *Symbiosis, 48*(1), 1-17.

Pèrez-Ramirez, N. O., Rogel, M. A., Wang, E., Castellanos, J. Z., & Martinez-Romero, E. (1998). Seeds of *Phaseolus vulgaris* bean carry *Rhizobium etli*. *FEMS Microbiology Ecology, 26*, 289-296.

Perret, X., Staehelin, C., & Broughton, W. J. (2000). Molecular basis of symbiotic promiscuity. *Microbiology and Molecular Biology Reviews, 64*(1), 180-201.

Pinto, F. G. S., Hungria, M., & Martins Mercante, F. (2007). Polyphasic characterization of Brazilian *Rhizobium tropici* strains effective in fixing N_2 with common bean (*Phaseolus vulgaris* L.). *Soil Biology and Biochemistry, 39*(8), 1851-1864.

Pitkäjärvi, J., Räsänen, L. A., Langenskiöld, J., Wallenius, K., Niemi, M., & Lindström, K. (2003). Persistence, population dynamics and competitiveness for nodulation of marker gene-tagged *Rhizobium galegae* strains in field lysimeters in the boreal climatic zone. *FEMS Microbiology Ecology, 46*(1), 91-104.

Poupot, R., Martinez-Romero, E., & Prome, J. C. (1993). Nodulation factors from *Rhizobium tropici* are sulfated or nonsulfated chitopentasaccharides containing an N-methyl-N-acylglucosaminyl terminus. *Biochemistry, 32*(39), 10430-10435.

Prell, J., Boesten, B., Poole, P., & Priefer, U. (2002). The *Rhizobium leguminosarum* bv. *viciae* VF39 aminobutyrate (GABA) aminotransferase gene (*gabT*) is induced by GABA and highly expressed in bacteroids. *Microbiology, 148*, 615-623.

Radutoiu, S., Madsen, L. H., Madsen, E. B., Felle, H. H., Umehara, Y., Gronlund, M., Sato, S., Nakamura, Y., Tabata, S., Sandal, N., & Stougaard, J. (2003). Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature, 425*(6958), 585-592.

Radutoiu, S., Madsen, L. H., Madsen, E. B., Jurkiewicz, A., Fukai, E., Quistgaard, E. M., Albrektsen, A. S., James, E. K., Thirup, S., & Stougaard, J. (2007). LysM domains mediate lipochitin–oligosaccharide recognition and Nfr genes extend the symbiotic host range. *The EMBO Journal, 26*(17), 3923-3935.

Rahmani, H. A., Räsänen, L. A., Afshari, M., & Lindström, K. (2011). Genetic diversity and symbiotic effectiveness of rhizobia isolated from root nodules of *Phaseolus vulgaris* L. grown in soils of Iran. *Applied Soil Ecology, 48*(3), 287-293.

Ramirez-Bahena, M. H., Garcia-Fraile, P., Peix, A., Valverde, A., Rivas, R., Igual, J. M., Mateos, P. F., Martinez-Molina, E., & Velazquez, E. (2008). Revision of the taxonomic status of the species *Rhizobium leguminosarum* (Frank 1879) Frank 1889^{AL}, *Rhizobium phaseoli* Dangeard 1926^{AL} and *Rhizobium trifolii* Dangeard 1926^{AL}. *R. trifolii* is a later synonym of *R. leguminosarum*. Reclassification of the strain *R. leguminosarum* DSM 30132 (=NCIMB 11478) as *Rhizobium pisi* sp. nov. *International Journal of Systematic and Evolutionary Microbiology, 58*(Pt 11), 2484-2490.

Ramírez, M., Flores-Pacheco, G., Reyes, J. L., Álvarez, A. L., Drevon, J. J., Girard, L., & Hernández, G. (2013). Two common bean genotypes with contrasting response to phosphorus deficiency show variations in the microRNA 399-mediated PvPHO2 regulation within the PvPHR1 signaling pathway. *International Journal of Molecular Sciences, 14*(4), 8328-8344.

Rao, J. R., Fenton, M., & Jarvis, B. D. W. (1994). Symbiotic plasmid transfer in *Rhizobium leguminosarum* biovar *trifolii* and competition between the inoculant strain icmp2163 and transconjugant soil bacteria. *Soil Biology and Biochemistry, 26*(3), 339-351.

Ratcliff, W. C., Kadam, S. V., & Denison, R. F. (2008). Poly-3-hydroxybutyrate (PHB) supports survival and reproduction in starving rhizobia. *FEMS Microbiology Ecology, 65*(3), 391-399.

Reeve, W. G., Tiwari, R. P., De Meyer, S. E., & Poole, P. (2016). Specialised genetic techniques for rhizobia. In J. G. Howieson & M. Dilworth (Eds.), *Working with rhizobia* (pp. 245-282). Canberra, Australia: Centre for International Agricultural Research.

Reeve, W. G., Tiwari, R. P., Worsley, P. S., Dilworth, M. J., Glenn, A. R., & Howieson, J. G. (1999). Constructs for insertional mutagenesis, transcriptional signal localization and gene regulation studies in root nodule and other bacteria. *Microbiology, 145*, 1307-1316.

Reid, D. E., Ferguson, B. J., & Gresshoff, P. M. (2011). Inoculation- and nitrate-induced CLE peptides of soybean control NARK-dependent nodule formation. *Molecular Plant-Microbe Interactions, 24*(5), 606-618.

Ribeiro, R. A., Martins, T. B., Ormeño-Orrillo, E., Delamuta, J. R. M., Rogel, M. A., Martínez-Romero, E., & Hungria, M. (2015). *Rhizobium ecuadorense* sp. nov., an indigenous N_2-fixing symbiont of the ecuadorian common bean (*Phaseolus vulgaris* L.) genetic pool. *International Journal of Systematic and Evolutionary Microbiology, 65*(9), 3162-3169.

Ribeiro, R. A., Rogel, M. A., Lopez-Lopez, A., Ormeno-Orrillo, E., Barcellos, F. G., Martinez, J., Thompson, F. L., Martinez-Romero, E., & Hungria, M. (2012). Reclassification of *Rhizobium tropici* type A strains as *Rhizobium leucaenae* sp. nov. *International Journal of Systematic and Evolutionary Microbiology, 62*(Pt 5), 1179-1184.

Richardson, A. E., Viccars, L. A., Watson, J. M., & Gibson, A. H. (1995). Differentiation of *Rhizobium* strains using the polymerase chain reaction with random and directed primers. *Soil Biology and Biochemistry, 27*, 515-524.

Rivas, R., Martens, M., de Lajudie, P., & Willems, A. (2009). Multilocus sequence analysis of the genus *Bradyrhizobium*. *Systematic and Applied Microbiology, 32*(2), 101-110.

Rodpothong, P., Sullivan, J. T., Songsrirote, K., Sumpton, D., Cheung, K. W. J. T., Thomas-Oates, J., Radutoiu, S., Stougaard, J., & Ronson, C. W. (2009). Nodulation gene mutants of *Mesorhizobium loti* R7A—*nodZ* and *nolL* mutants have host-specific phenotypes on *Lotus* spp. *Molecular Plant-Microbe Interactions, 22*(12), 1546-1554.

Rogel, M. A., Ormeño-Orrillo, E., & Martinez Romero, E. (2011). Symbiovars in rhizobia reflect bacterial adaptation to legumes. *Systematic and Applied Microbiology, 34*(2), 96-104.

Román-Ponce, B., Jing Zhang, Y., Soledad Vásquez-Murrieta, M., Hua Sui, X., Feng Chen, W., Carlos Alberto Padilla, J., Wu Guo, X., Lian Gao, J., Yan, J., Hong Wei, G., & Tao Wang, E. (2016). *Rhizobium acidisoli* sp. nov., isolated from root nodules of *Phaseolus vulgaris* in acid soils. *International Journal of Systematic and Evolutionary Microbiology, 66*(1), 398-406.

Rosselló-Móra, R., & Amann, R. (2001). The species concept for prokaryotes. *FEMS Microbiology Reviews, 25*, 39-67.

Rouhrazi, K., Khodakaramian, G., & Velázquez, E. (2016). Phylogenetic diversity of rhizobial species and symbiovars nodulating *Phaseolus vulgaris* in Iran. *FEMS Microbiology Letters, 363*(5).

Sadowsky, M. J., & Cregan, P. B. (1992). The soybean Rj4 allele restricts nodulation by *Bradyrhizobium japonicum* serogroup 123 strains. *Applied and Environmental Microbiology, 58*(2), 720-723.

Šalplachta, J., Kubesová, A., Moravcová, D., Vykydalová, M., Süle, S., Matoušková, H., Horký, J., & Horká, M. (2013). Use of electrophoretic techniques and MALDI–TOF MS for rapid and reliable characterization of bacteria: analysis of intact cells, cell lysates, and "washed pellets". *Analytical and Bioanalytical Chemistry, 405*(10), 3165-3175.

Sanchez-Canizares, C., & Palacios, J. (2013). Construction of a marker system for the evaluation of competitiveness for legume nodulation in *Rhizobium* strains. *Journal of Microbiological Methods, 92*(3), 246-249.

Sarita, S., Sharma, P. K., Priefer, U. B., & Prell, J. (2005). Direct amplification of rhizobial *nodC* sequences from soil total DNA and comparison to *nodC* diversity of root nodule isolates. *FEMS Microbiology Ecology, 54*(1), 1-11.

Schmutz, J., McClean, P. E., Mamidi, S., Wu, G. A., Cannon, S. B., Grimwood, J., Jenkins, J., Shu, S., Song, Q., Chavarro, C., Torres-Torres, M., Geffroy, V., Moghaddam, S. M., Gao, D., Abernathy, B., Barry, K., Blair, M., Brick, M. A., Chovatia, M., Gepts, P., Goodstein, D. M., Gonzales, M., Hellsten, U., Hyten, D. L., Jia, G., Kelly, J. D., Kudrna, D., Lee, R., Richard, M. M. S., Miklas, P. N., Osorno, J. M., Rodrigues, J., Thareau, V., Urrea, C. A., Wang, M., Yu, Y., Zhang, M., Wing, R. A., Cregan, P. B., Rokhsar, D. S., & Jackson, S. A. (2014). A reference genome for common bean and genome-wide analysis of dual domestications. *Nature Genetics, 46*(7), 707-713.

Seefeldt, L. C., Dance, I. G., & Dean, D. R. (2004). Substrate interactions with nitrogenase: Fe versus Mo. *Biochemistry, 43*(6), 1401-1409.

Segovia, L., Young, J. P. W., & Martinez-Romero, E. (1993). Reclassification of American *Rhizobium leguminosarum* biovar *phaseoli* Type I strains as *Rhizobium etli* sp. nov. *International Journal of Systematic and Evolutionary Microbiology, 43*(2), 374-377.

Sessitsch, A., Hardarson, G., de Vos, W., & Wilson, K. (1998). Use of marker genes in competition studies of *Rhizobium*. *Plant and Soil, 204*, 35-45.

Sessitsch, A., Wilson, K., Akkermans, A., & de Vos, W. (1996). Simultaneous detection of different *Rhizobium* strains marked with either the *Escherichia coli gusA* gene or the *Pyrococcus furiosus celB* gene. *Applied and Environmental Microbiology 62*(11), 4191-4194.

Shamseldin, A. (2007). Use of DNA marker to select well-adapted *Phaseolus*-symbionts strains under acid conditions and high temperature. *Biotechnology Letters, 29*(1), 37-44.

Shaner, N. C., Lambert, G. G., Chammas, A., Ni, Y., Cranfill, P. J., Baird, M. A., Sell, B. R., Allen, J. R., Day, R. N., Israelsson, M., Davidson, M. W., & Wang, J. (2013). A bright monomeric green fluorescent protein derived from *Branchiostoma lanceolatum*. *Nat Meth, 10*(5), 407-409.

Silva, F., Queiroz, J. A., & Domingues, F. (2011). Evaluating metabolic stress and plasmid stability in plasmid DNA production by *Escherichia coli*. *Biotechnology Advances, 30*, 691-708.

Singleton, P. W., & Stockinger, K. R. (1983). Compensation against ineffective nodulation in soybean. *Crop Science, 23*(1), 69-72.

Singleton, P. W., & Tavares, J. W. (1986). Inoculation response of legumes in relation to the number and effectiveness of indigenous *Rhizobium* populations. *Applied and Environmental Microbiology, 51*(5), 1013-1018.

Smil, V. (2001). *Enriching the earth: Fritz Haber, Carl Bosch, and the transformation of world food production.* : MIT Press.

Sobecky, P. A., Easter, C. L., Bear, P. D., & Helinski, D. R. (1996). Characterization of the stable maintenance properties of the *par* region of the broad-host-range plasmid RK2. *Journal of Bacteriology, 178*(7), 2086-2093.

Spaink, H. P., Wijffelman, C. A., Pees, E., Okker, R. J. H., & Lugtenberg, B. J. J. (1987). *Rhizobium* nodulation gene *nodD* as a determinant of host specificity. *Nature, 328*(6128), 337-340.

Sprent, J. I. (2007). Evolving ideas of legume evolution and diversity: a taxonomic perspective on the occurrence of nodulation. *The New Phytologist, 174*(1), 11-25.

Sprent, J. I. (2008). 60Ma of legume nodulation. What's new? What's changing? *Journal of Experimental Botany, 59*(5), 1081-1084.

Sprent, J. I., Ardley, J. K., & James, E. K. (2013). From North to South: A latitudinal look at legume nodulation processes. *South African Journal of Botany, 89*, 31-41.

Spriggs, A. C., & Dakora, F. D. (2009). Assessing the suitability of antibiotic resistance markers and the indirect ELISA technique for studying the competitive ability of selected *Cyclopia* Vent. rhizobia under glasshouse and field conditions in South Africa. *BMC Microbiology, 9*, 142-142.

Spurr, A. R. (1969). A low-viscosity epoxy resin embedding medium for electron microscopy. *Journal of Ultrastructure Research, 26*, 31-43.

Stackebrandt, E., & Goebel, B. M. (1994). Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *International Journal of Systematic and Evolutionary Microbiology, 44*(4), 846-849.

Streit, W., Koch, K., & Werner, D. (1992). Nodulation competitiveness of *Rhizobium leguminosarum* bv. *phaseoli* and *Rhizobium tropici* strains measured by glucuronidase (*gus*) gene fusion. *Biology and Fertility of Soils, 14*, 140-144.

Sullivan, J. T., Patrick, H. N., Lowther, W. L., Scott, D. B., & Ronson, C. W. (1995). Nodulating strains of *Rhizobium loti* arise through chromosomal symbiotic gene transfer in the environment. *Proceedings of the National Academy of Sciences, 92*(19), 8985-8989.

Talbi, C., Delgado, M. J., Girard, L., Ramirez-Trujillo, A., Caballero-Mellado, J., & Bedmar, E. J. (2010). *Burkholderia phymatum* strains capable of nodulating *Phaseolus vulgaris* are present in Moroccan soils. *Applied and Environmental Microbiology, 76*(13), 4587-4591.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: Molecular evolutionary genetics analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony methods. *Molecular Biology and Evolution, 28*(10), 2731-2739.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution, 30*, 2725-2729.

Tate, R., Patriarca, E. J., Riccio, A., Defez, R., & Iaccarino, M. (1994). Development of *Phaseolus vulgaris* root nodules. *Molecular Plant-Microbe Interactions, 7*(5), 582-589.

Terpolilli, J. J. (2009). *Why are the symbiosis between some genotypes of Sinorhizobium and Medicago suboptimal for N_2 fixation?* (Doctor of Philosophy), Murdoch University, Perth, Australia.

Terpolilli, J. J., Masakapalli, S. K., Karunakaran, R., Webb, I., Green, R., Watmough, N. J., Kruger, N. J., Ratcliffe, R. G., & Poole, P. S. (2016). Lipogenesis and redox balance in nitrogen-fixing pea bacteroids. *Journal of Bacteriology*.

Thies, J., Singleton, P. W., & Bohlool, B. (1991a). Influence of the size of indigenous rhizobial populations on establishment and symbiotic performance of introduced rhizobia on field-grown legumes. *Applied and Environmental Microbiology, 57*(1), 19-28.

Thies, J., Singleton, P. W., & Bohlool, B. (1991b). Modeling symbiotic performance of introduced rhizobia in the field by use of indices of indigenous population size and nitrogen status of the soil. *Applied and Environmental Microbiology, 57*(1), 29-37.

Thies, J. E., Holmes, E. M., & Vachot, A. (2001). Application of molecular techniques to studies in *Rhizobium* ecology: a review. *Australian Journal of Experimental Agriculture, 41*(3), 299-319.

Thoma, S., & Schobert, M. (2009). An improved *Escherichia coli* donor strain for diparental mating. *FEMS Microbiology Letters, 294*(2), 127-132.

Tilman, D., Fargione, J., Wolff, B., D'Antonio, C., Dobson, A., Howarth, R., Schindler, D., Schlesinger, H. W., Simberloff, D., & Swackhamer, D. (2001). Forecasting agriculturally driven global environmental change. *Science, 292*(5515), 281-284.

Tindall, B. J., Rosselló-Móra, R., Busse, H.-J., Ludwig, W., & Kämpfer, P. (2010). Notes on the characterization of prokaryote strains for taxonomic purposes. *International Journal of Systematic and Evolutionary Microbiology, 60*(1), 249-266.

Trainer, M. A. (2009). *Carbon metabolism and desiccation tolerance in the nitrogen-fixing rhizobia Bradyrhizobium japonicum and Sinorhizobium meliloti.* (PhD), Waterloo, Ontario Canada.

Trainer, M. A., & Charles, T. C. (2006). The role of PHB metabolism in the symbiosis of rhizobia with legumes. *Applied Microbiology and Biotechnology, 71*(4), 377-386.

Triplett, E. W., & Sadowsky, M. J. (1992). Genetics of competition for nodulation of legumes. *Annual Review of Microbiology, 46*(1), 399-422.

Tsai, S. M., Bonetti, R., Agbala, S. M., & Rossetto, R. (1993). Minimizing the effect of mineral nitrogen on biological nitrogen fixation in common bean by increasing nutrient levels. *Plant and Soil, 152*(1), 131-138.

Tsukui, T., Eda, S., Kaneko, T., Sato, S., Okazaki, S., Kakizaki-Chiba, K., Itakura, M., Mitsui, H., Yamashita, A., Terasawa, K., & Minamisawa, K. (2013). The Type III secretion system of *Bradyrhizobium japonicum* USDA122 mediates symbiotic incompatibility with Rj2 soybean plants. *Applied and Environmental Microbiology, 79*(3), 1048-1051.

Udvardi, M., & Poole, P. S. (2013). Transport and metabolism in legume-rhizobia symbioses. *Annual Review of Plant Biology, 64*, 781-805.

Vadez, V., Lasso, J. H., Beck, D. P., & Drevon, J. J. (1999). Variability of N_2-fixation in common bean (*Phaseolus vulgaris* L.) under P deficiency is related to P use efficiency. *Euphytica, 106*(3), 231-242.

Valverde, A., Igual, J. M., Peix, A., Cervantes, E., & Velazquez, E. (2006). *Rhizobium lusitanum* sp. nov. a bacterium that nodulates *Phaseolus vulgaris*. *International Journal of Systematic and Evolutionary Microbiology, 56*(Pt 11), 2631-2637.

Van Berkum, P., Beyene, D., Bao, G., Campbell, A., & Eardly, B. (1998). *Rhizobium mongolense* sp. nov is one of three rhizobial genotypes identified which nodulate and form nitrogen-fixing symbioses with *Medicago ruthenica* [(L) Ledebour]. *International Journal of Systematic Bacteriology, 48*, 13-22.

van der Bom, F. (2012). *Response of Phaseolus vulgaris L. and Pisum sativum L. to N and P fertilisers and inoculation with Rhizobium in Kennya.* (Master of Science), Wageningen University.

Vanlauwe, B., Coe, R., & Giller, K. E. (2016). Beyond averages: new approaches to understand heterogeneity and risk of technology success or failure in smallholder farming. *Experimental Agriculture*, 1-23.

Vargas, T. M. A., Mendes, C. I., & Hungria, M. (2000). Response of field-grown bean (*Phaseolus vulgaris* L.) to Rhizobium inoculation and nitrogen fertilization in two Cerrados soils. *Biology and Fertility of Soils, 32*(3), 228-233.

Venable, J. H., & Coggeshall, R. (1965). A simplified lead citrate stain for use in electron microscopy. *The Journal of Cell Biology, 25*(2), 407-408.

Vinuesa, P., Silva, C., Lorite, M. J., Izaguirre-Mayoral, M. L., Bedmar, E. J., & Martínez-Romero, E. (2005). Molecular systematics of rhizobia based on maximum likelihood and Bayesian phylogenies inferred from *rrs, atpD, recA* and *nifH* sequences, and their use in the classification of *Sesbania* microsymbionts from Venezuelan wetlands. *Systematic and Applied Microbiology, 28*(8), 702-716.

Vitousek, P. M., Menge, D. N. L., Reed, S. C., & Cleveland, C. C. (2013). Biological nitrogen fixation: rates, patterns and ecological controls in terrestrial ecosystems. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 368*(1621).

Vitousek, P. M., Mooney, H. A., Lubchenco, J., & Melillo, J. M. (1997). Human domination of earth's ecosystems. *Science, 277*(5325), 494-499.

Vlassak, K., Vanderleyden, J., & Franco, A. (1996). Competition and persistence of *Rhizobium tropici* and *Rhizobium etli* in tropical soil during successive bean (*Phaseolus vulgaris* L.) cultures. *Biology and Fertility of Soils, 21*(1), 61-68.

Vlassak, K. M., Vanderleyden, J., & Graham, P. H. (1997). Factors influencing nodule occupancy by inoculant rhizobia. *Critical Reviews in Plant Sciences, 16*(2), 163-229.

Voorhorst, W. G., Eggen, R. I., Luesink, E. J., & de Vos, W. M. (1995). Characterization of the *celB* gene coding for beta-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus* and its expression and site-directed mutation in *Escherichia coli*. *Journal of Bacteriology, 177*(24), 7105-7111.

Walker, S. A., & Downie, J. A. (2000). Entry of *Rhizobium leguminosarum* bv. *viciae* into root hairs requires minimal nod factor specificity, but subsequent infection thread growth requires *nodO* or *nodE*. *Molecular Plant-Microbe Interactions, 13*(7), 754-762.

Wang, F., Wang, E. T., Wu, L. J., Sui, X. H., Li, Y., & Chen, W. X. (2011). *Rhizobium vallis* sp. nov., isolated from nodules of three leguminous species. *International Journal of Systematic and Evolutionary Microbiology, 61*(11), 2582-2588.

Wang, Y., Zhang, Z., & Ramanan, N. (1997). The actinomycete Thermobispora bispora contains two distinct types of transcriptionally active 16S rRNA genes. *Journal of Bacteriology, 179*(10), 3270-3276.

Waswa, M. N., Karanja, N. K., Woomer, P. L., & Mwenda, G. M. (2014). Identifying elite rhizobia for soybean (*Glycine max*) in Kenya. *African Journal of Crop Science, 2*(2), 60-66.

Weaver, R. W., & Frederick, L. R. (1974). Effect of inoculum rate on competitive nodulation of *Glycine max* L. Merrill. II. Field studies. *Agronomy Journal, 66*(2), 233-236.

Weilbo, J., Kuske, J., Marek-Kozaczuk, M., & Skorupska, A. (2010). The competition between *Rhizobium leguminosarum* bv. *viciae* strains progresses until late stages of symbiosis. *Plant and Soil, 337*, 125-135.

Weinstein, M., Roberts, R. C., & Helinski, D. R. (1992). A region of the broad-host-range plasmid RK2 causes stable in planta inheritance of plasmids in *Rhizobium meliloti* cells isolated from alfalfa root nodules. *Journal of Bacteriology, 174*(22), 7486-7489.

White, J., Prell, J., James, E. K., & Poole, P. (2007). Nutrient sharing between symbionts. *Plant Physiology, 144*(2), 604-614.

Wiedenbeck, J., & Cohan, F. M. (2011). Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. *FEMS Microbiology Reviews, 35*(5), 957-976.

Wielbo, J., Marek-Kozaczuk, M., Kubik-Komar, A., & Skorupska, A. (2007). Increased metabolic potential of *Rhizobium* spp. is associated with bacterial competitiveness. *Canadian Journal of Microbiology, 53*(8), 957-967.

Williams, A., Wilkinson, A., Krehenbrink, M., Russo, D. M., Zorreguieta, A., & Downie, J. A. (2008). Glucomannan-mediated attachment of *Rhizobium leguminosarum* to pea root hairs is required for competitive nodule infection. *Journal of Bacteriology, 190*(13), 4706-4715.

Willis, L. B., & Walker, G. C. (1998). The *phbC* (poly-β-hydroxybutyrate synthase) gene of *Rhizobium* (*Sinorhizobium*) *meliloti* and characterization of *phbC* mutants. *Canadian Journal of Microbiology, 44*(6), 554-564.

Wilson, K. J., Sessitsch, A., Corbo, J. C., Giller, K., Akkermans, A., & Jefferson, R. A. (1995). B-Glucuronidase(GUS) transposons for ecological and genetic studies of rhizobia and other Gram-negativebacteria. *Microbiology, 141*, 1691-1705.

Witty, J. F., Minchin, F. R., & Sheehy, J. E. (1983). Carbon costs of nitrogenase activity in legume root nodules determined using acetylene and oxygen. *Journal of Experimental Botany, 34*(145), 951-963.

Woese, C. R. (1987). Bacterial evolution. *Microbiological Reviews, 51*(2), 221-271.

Wongphatcharachai, M., Wang, P., Staley, C., Chun, C. L., Ferguson, J. A., Moncada, K. M., Sheaffer, C. C., & Sadowsky, M. J. (2015). Site-specific distribution and competitive ability of indigenous bean-nodulating rhizobia isolated from organic fields in Minnesota. *Journal of Biotechnology, 214*, 158-168.

Woomer, P. (1990). *Predicting the abundance of indigenous and the persistence of introduced rhizobia in tropical soils.* (PhD), University of Hawaii, Hawaii.

Yano, H., Deckert, G. E., Rogers, L. M., & Top, E. M. (2012). Roles of long and short replication initiation proteins in the fate of IncP-1 plasmids. *Journal of Bacteriology, 194*(6), 1533-1543.

Yarza, P., Yilmaz, P., Pruesse, E., Glockner, F. O., Ludwig, W., Schleifer, K.-H., Whitman, W. B., Euzeby, J., Amann, R., & Rossello-Mora, R. (2014). Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nature Reviews Microbiology, 12*(9), 635-645.

Yates, Howieson, J. G., Hungria, M., Bala, A., O'Hara, G., & Terpolilli, J. J. (2016a). Authentication of rhizobia and assessment of the legume symbiosis in controlled growth systems. In J. G. Howieson & M. Dilworth (Eds.), *Working with rhizobia* (pp. 73-108). Canberra, Australia: Australian Centre for International Agricultural Research.

Yates, R. J., Abaidoo, R., & Howieson, J. G. (2016b). Field experiments with rhizobia. In J. G. Howieson & M. Dilworth (Eds.), *Working with rhizobia* (pp. 145-166). Canberra, Australia: Australian Centre for International Agricultural Research.

Yates, R. J., Howieson, J. G., Reeve, W. G., Brau, L., Speijers, J., Nandasena, K., Real, D., Sezmis, E., & O'Hara, G. W. (2008). Host–strain mediated selection for an effective nitrogen-fixing symbiosis between *Trifolium* spp. and *Rhizobium leguminosarum* biovar *trifolii*. *Soil Biology and Biochemistry, 40*(3), 822-833.

Yates, R. J., Howieson, J. G., Reeve, W. G., & O'Hara, G. W. (2011). A re-appraisal of the biology and terminology describing rhizobial strain success in nodule occupancy of legumes in agriculture. *Plant and Soil, 348*(1-2), 255-267.

Zahran, H. H. (1999). *Rhizobium*-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. *Microbiology and Molecular Biology Reviews, 63*(4), 968-989.

Zevenhuizen, L. P. T. M. (1981). Cellular glycogen β-1,2-glucan, poly-β-hydroxytutyric acid and extracellular polysaccharides in fast-growing *Rhizobium* species. *Antonie van Leeuwenhoek, 47*, 481-497.

Zurdo-Piñeiro, J. L., García-Fraile, P., Rivas, R., Peix, A., León-Barrios, M., Willems, A., Mateos, P. F., Martínez-Molina, E., Velázquez, E., & van Berkum, P. (2009). Rhizobia from Lanzarote, the Canary Islands, that nodulate *Phaseolus vulgaris* have characteristics in common with *Sinorhizobium meliloti* isolates from mainland Spain. *Applied and Environmental Microbiology, 75*(8), 2354-2359.