

**DIVERSITY AND DISTRIBUTION OF *Tylosema*
esculentum (MARAMA BEAN) ENDOPHYTIC
BACTERIA COMMUNITIES IN OMITARA,
HARNAS AND OTJINENE, EASTERN NAMIBIA**

A DISSERTATION SUBMITTED IN FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY (MICROBIOLOGY)

OF

THE UNIVERSITY OF NAMIBIA

BY

Jean Damascène UZABAKIRIHO

200535633

March 2016

Supervisor: Prof. PM Chimwamurombe

ABSTRACT

Tylosema esculentum is a nutritious drought avoiding plant endemic to the Kalahari Desert. Our study assessed the density, diversity and distribution of endophytic microbial community structures associated with leaves, stems and tuberous roots of *T. esculentum* in Eastern Namibia. Culture-dependent and PCR-based 454 pyrosequencing methods were used. ANOVA with pairwise comparison revealed a significant difference in bacterial density between below and above ground. Endophytic bacterial isolates (605) were identified, grouped into 24 genera and three phyla. *Proteobacteria* was the most represented (67.4%) followed by *Firmicutes* (23.7%) and *Actinobacteria* (4.3%). Shannon diversity index, revealed a significant difference between the tuberous roots and leaves ($p = 0.005$) and stems ($p = 0.006$) microbial communities. The cluster analysis revealed a separation between the above and below ground microbial communities. The PCA and the Jaccard diversity indices confirmed these findings. Our results suggested that the microbial community composition was mainly governed by the plant parts rather than the location or sampling time. The phylogenetic analysis showed that all these microbial communities fell into two clades distinct from known cultivated bacteria from NCBI. All isolates associated with *T. esculentum* were positive for the *nifH* gene amplification. Only 42% nested with known strain in the NCBI GeneBank Database. This finding showed the presence of putative novel nitrogen-fixing bacteria associated with *T. esculentum*. Ten samples (leaves, stems and tuberous roots) from Omitara were examined using 16S rRNA gene pyrosequencing method. The presence of the three phyla *Firmicutes* (50.3%), *Proteobacteria* (38.32%) and *Actinobacteria* (4.46%) was confirmed. Two more phyla *Fusobacteria* (5.7%) and *Bacteroidetes* (1.368%) were revealed. Strikingly, 2 phyla (*Firmicutes* and *Proteobacteria*) of the five phyla represented 89 % of the total sequences. Similarly, four families (*Enterobacteriaceae*, *Bacillaceae*, *Pasteurellaceae* and *Fusobacteriaceae*) of the 18 recorded, represented nearly 91% of the total *T. esculentum* bacterial community assemblage. The genus *Bacillus* was predominantly found in the tuberous roots though shared across all samples. The class *Clostridiales* was exclusively found in leaves and stems. Within the *Gammaproteobacteria* class, the sequence grouping showed the *Enterobacteriaceae* family and the yet to be identified *Enterobacteriaceae* dominated. Unlike the *Enterobacteriaceae* that are mostly found in the tuberous roots, the *Pasteurellaceae* family was preferably found in the leaves and stems. *Actinobacteria* have shown a ubiquitous colonization compared with the *Bacteroidetes* that colonized the above ground part of the plant. *T. esculentum* organs exerted selective pressures on their associated bacterial communities. Only 68% of all reads could be classified at the phylum level. *Firmicutes* are the most dominant phylum in the current study with 46.9% sequences that have not yet been classified to any existing family, order or genus. Also, the rarefaction curves predicted that additional sampling will lead to significantly increased estimates of diversity. Sequences in this study, have shown similarities with sequences occurring in water-stressed environments with plant growth promoting traits. In conclusion, *T. esculentum* bean lives in community with a large diversity of potentially plant growth promoting bacteria

DEDICATION

To my father Thaddée UZABAKIRIHO and my father-in-law Marc BIZIMANA upon whose shoulders I stand;

To all other family members and nameless tens of thousands of innocent people who perished in wars of conquest and tyranny in the African Great Lakes Region; whose respect and compassion has been downplayed to insult their memory;

To my mothers Verdiane HITIMANA and Bernadette MUKESHIMANA for your indomitable spirit and your incredible strength, against all odds;

To my brilliant and loving wife, Brigitte NSHIMYIMANA who has accompanied me with unlimited patience, understanding, help and encouragement; To my high-spirited, kind-hearted and industrious son, Jean Emmery NGABO-SHINGIRO and our sedulous and assiduous niece Flora Tuyisenge NIYONZIMA,

This thesis is dedicated

ACKNOWLEDGEMENT

Voltaire, French author and humanist (1694 - 1778) once said “Appreciation is a wonderful thing, it makes what is excellent in others belong to us as well”. And bestowing to an old dictum: “If you never learn the language of gratitude, you’ll never be on speaking terms with happiness”. And I agree. When we succeed, we do so because others have abetted.

I am extremely indebted to my supervisor Prof. Percy M. Chimwamurombe to whom I would like to express my deepest sense of gratitude. This work would not have been possible without his valuable guidance, support, patience and encouragement. He provided directions to recover when my steps wavered. He found solutions where others would only see problems. I will carry with me the lessons gained as I pursue this rewarding career.

I received generous support from the Head of Department of Biological Sciences Dr. Elsabe Julies and Staff members who gave me access to the laboratory and research facilities. Without your precious contribution, it would have been impossible to conduct this research. Discussions with Dr. Ezekiel Kwembeya, Ms. Fransiska Kangombe and Dr. Seth Eiseb have been illuminating with their thought-provoking contributions. Thanks for your constructive comments and warm encouragement. I have had the support and inspiration of Prof. Jean Baptiste Gatsinzi and Mr. Dismas Ntirampeba. As brothers, you were always more than willing to help and give your best suggestions in Programming and Statistics at the time where it was the most needed. I greatly appreciate your excellent

assistance and spiritual support. I am thankful to my lab mates, Dr. D. Kavishe, Dr E. Nepolo and Dr. M. Takundwa for the valuable late hours spent together. I am grateful for the assistance given by Mr. Mbangu, Ms. Kaitjizemine, Mrs. Morkel, Ms. Johnson, Mr. Daniel Haiyambo, Mr. Billy Chinsenga and Ms. Celine Mukakalisa. Your care helped me to overcome setbacks. Lastly, and most importantly my family, none of these would have happened without your love, patience and profound understanding. You were always there cheering me up and stood by me through the good and bad times.

Soli Deo Gloria

DECLARATION

I, Jean Damascène UZABAKIRIHO, declare hereby that this study is a true reflection of my own research, and that this work, or part thereof has not been submitted for a degree in any other institution of higher education. .

No part of this thesis/dissertation may be reproduced, stored in any retrieval system, or transmitted in any form, or by means (e.g. electronic, mechanical, photocopying, recording or otherwise) without the prior permission of the author, or The University of Namibia in that behalf. .

I, Jean Damascène UZABAKIRIHO, grant The University of Namibia the right to reproduce this thesis in whole or in part, in any manner or format, which The University of Namibia may deem fit, for any person or institution requiring it for study and research; providing that The University of Namibia shall waive this right if the whole thesis has been or is being published in a manner satisfactory to the University.

Jean Damascène UZABAKIRIHO

Date

ACRONYMS

DNA/RNA	Deoxyribonucleic acid/ Ribonucleic acid
NGS	Next Generation Sequencing
HTS	High-Throughput Sequencing
BLAST	Basic Local Alignment Search Tool
RDP	Ribosomal Database Project
OTU	Operational Taxonomic Unit
PCR	Polymerase Chain Reaction
DGGE/TGGE	Denaturing Gradient/Temperature Gradient Gel Electrophoresis
T-RFLP	Terminal Restriction Fragment Length Polymorphism Analysis
PGPB	Plant Growth Promoting Bacteria
PCA	Principal Component Analysis
IAA	Indole-3-acetic acid
TSA	Tryptic Soy Agar
SIM	Synthetic Malate

LIST OF TABLE

Table 1: Enumeration of putative endophytic diazotroph bacteria isolated from aseptically fresh <i>T. esculentum</i> tissue leaves (L), stems (S) and tuberous roots (T) (means of three replicates, $\pm$ standard deviation) from different locations (Omitara, Harnas and Otjinene) over a period of 4 years (2011-2014).-----	54
Table 2: Two-way ANOVA showing that there is enough evidence to indicate that at least one of the three plant parts has a mean CFU different from other parts.-----	55
Table 3: Two-way ANOVA indicating that at 5% significance level at least one of the three plant parts has a mean CFU different from other parts (as p-value=0.000 is less than 0.05). -----	56
Table 4: Results of multiple comparisons performed on the total amount of bacterial cells to determine -----	57
Table 5: Phylum and class distribution frequencies (%) of endophytic bacteria associated with <i>T. esculentum</i> 's tissues (leaves, stems and tuberous roots). -----	59
Table 6: Two-way ANOVA of the effect on the presence of bacterial phyla occurring in different <i>Tylosema esculentum</i> 's plant parts (leaf, stem and tuberous root). -----	61
Table 7: Two-way ANOVA of the effect on the presence of bacterial phyla occurring in different <i>Tylosema esculentum</i> 's plant parts (leaf, stem and tuberous root). -----	66
Table 8: Test for equal means for dominance index -----	67
Table 9: Comparison of culturable sequences associated with <i>T. esculentum</i> -----	81

Table 10: Sequence similarity between diazotrophic bacterial isolates and the closest type strain from the NCBI Database. A Comparison between the first hit and cultured bacteria. -----	196
Table 11: Sequence similarity between diazotrophic bacterial isolates and the closest type strain from the NCBI Database. A Comparison between the first hit and cultured bacteria (Continued). -----	197
Table 12: Eigenvalues and percentage of explained variability extracted from <i>T. esclentum</i> sequences based on 16S rRNA gene sequences generated using Roche/454 pyrosequencing. -----	198

LIST OF FIGURES

- Figure 1: Distribution map of the genus *Tylosema*: *T. angolense* (○), *T. argenteum* (+), *T. esculentum* (◆), *T. fassoglense* (●), *T. humifusum* (Δ, Δ), unidentified specimens (○) adapted from (Castro et al., 2005). -----8
- Figure 2: *T. esculentum* (Marama bean) plant with its tuberous root in Harnas (Omaheke region, Eastern Namibia) (A) and *T. esculentum* mature dry seeds (B)-----9
- Figure 3: Hierarchical cluster dendrogram of bacterial communities' assemblages, plant parts and their respective locations. The sample designation shows plant parts by L, S, T as leaf, stem and tuberous root respectively; OM for Omitara, HARN for Harnas and OTJ for Otjinene and lastly the sampling year.----- 63
- Figure 4: Boxplot that compared Shannon's diversity indices between the three different plants parts (leaves, stem and tuberous roots). Minimum and maximum values are indicated by the bars above boxes with the median values indicated by the inner blue box----- 66
- Figure 5: Boxplot that compared Simpson abundance values between the three different plant parts (leaves, stems and tuberous root). Minimum and maximum values are indicated by the bars above boxes with the median values indicated by the inner blue box.----- 66
- Figure 6: Loading of PC1 of community bacterial samples (microbial genus were used as variables in PCA) ----- 71
- Figure 7: Loadings of PC2 of bacterial community bacterial samples (microbial genus were used as variables in PCA). ----- 72
- Figure 8: Clustering relationships between endophytic microbial communities based on principal component analysis (PCA). Here are shown figure (A): the total grouping of endophytic microbial communities in *T. esculentum*; (B), (C) and (D) separately clustered leaves, stems and tuberous root have shown a clear separation.----- 76

Figure 9: Strict consensus tree showing the phylogenetic relatedness of recovered endophytic bacteria associated with *T. esculentum* based on 16S rRNA sequences. The phylogenetic tree represents a maximum parsimony analysis of 143 taxa. Bootstrap values higher or equal to 50% (500 replicates) are shown at each branch. ----- 78

Figure 10: Strict consensus of the most parsimonious tree inferred using 16S rRNA sequences associated with *T. esculentum* based on 16S rRNA sequences. The phylogenetic tree represents a maximum parsimony analysis of 143 taxa with their reference strain. Bootstrap values higher than or equal to 50% (500 replicates) are shown at each branches.----- 80

Figure 11: Maximum parsimony tree of all *nifH* gene sequence similarity, showing the position of nitrogen fixing isolates obtained from *T. esculentum*. Numbers at the nodes represent the level of bootstrap support based on 500 replicates, only values > 50% were indicated. ----- 88

Figure 12: Maximum parsimony tree of *nifH* gene sequence similarity, showing the position of nitrogen fixing isolates obtained from *T. esculentum* and their closest similar match from the NCBI database. Numbers, at the nodes show the level of bootstrap support based on 500 replicates, only values >50% were represented. -- 89

Figure 13: *NifH* distribution in 40 selected bacterial samples----- 90

Figure 14: Bacteria phyla distribution recovered from *T. esculentum* using 454 sequencing methods. ----- 92

Figure 15: Above ground phyla distribution in *T. esculentum* as revealed by 454 sequencing approach.----- 93

Figure 16: Below ground phyla distribution in *T. esculentum* revealed by 454 sequencing approach. ----- 93

Figure 18: UPGMA dendrogram of 10 *T. esculentum* samples based on 16S rRNA gene sequences generated using Roche"454 pyrosequencing----- 98

- Figure 19: PCA ordination diagram illustrating the relationship among plots of PC1 versus PC2 scores of *T. esculentum* microbial communities distinguishing samples above (leaves and stems) and belowground (tuberous tuber). ----- 100
- Figure 20: Loading plot for PC1 for total *T. esculentum* microbial community based on 16S rRNA gene sequences generated using Roche/454 pyrosequencing. ----- 103
- Figure 21: Loading plot for PC 2 for total *T. esculentum* microbial community based on 16S rRNA gene sequences generated using Roche/454 pyrosequencing. ----- 104
- Figure 22: Loading plot for PC1 for leaves and stems *T. esculentum* microbial community. ----- 105
- Figure 23: Loading plot for PC2 for leaves and stems *T. esculentum* microbial community. ----- 106
- Figure 24: Rarefaction curves of *T. esculentum* microbial community based on 16S rRNA gene sequences generated using Roche /454 pyrosequencing (A) family and (B) genus at 0.03 distances. ----- 107
- Figure 24 (Cont'd): Rarefaction curves of *T. esculentum* microbial community based on 16S rRNA gene sequences generated using Roche/454 pyrosequencing (A) Family and (B) genus at a 0.03 distance. ----- 108

TABLE OF CONTENTS

ABSTRACT	II
DEDICATION.....	III
ACKNOWLEDGEMENT	IV
DECLARATION.....	VI
ACRONYMS	VII
LIST OF TABLE.....	VIII
LIST OF FIGURES	X
TABLE OF CONTENTS.....	XIII
CHAPTER 1: INTRODUCTION	1
1.1 Background of the study.....	1
1.2 Statement of the research problem.....	4
1.3 Objectives of the study	5
1.4 Significance of the study	6
CHAPTER 2: LITERATURE REVIEW	7
2.1 <i>Tylosema esculentum</i> (Marama bean): A plant with great agricultural potential.	7
2.1.1 Botanical and ecological description of <i>T. esculentum</i>	7
2.1.2 <i>T. esculentum</i> 's nutritional value.....	10
2.1.3 Uses and potential uses of <i>T. esculentum</i>	12

2.2 Endophytic microorganisms.....	13
2.2.1 Mechanisms stabilizing plant-microbe symbiosis.....	14
2.2.2 Vertical or transovarian transmission.....	14
2.2.3 Horizontal or environmental microbial transmission.....	15
2.3 Plant microbial interaction and resource partitioning.....	16
2.3.1 Plant – microbe interaction.....	16
2.3.2 Microbial niche partitioning: a response to resource availability.....	18
2.3.3 Plant microbial communities.....	19
2.3.4 Plant endophytic microbial communities: Assemblage and diversity.....	22
2.3.5 Plant Microbial communities fine-tuning.....	25
2.3.6 Benefits of microbial community diversity analysis.....	26
2.4 Plant Microbial Diversity Analysis.....	27
2.4.1 Denaturing Gradient Gel Electrophoresis (DGGE) and Temperature Gradient Gel Electrophoresis (TGGE).....	30
2.4.2 Terminal Restriction Fragment Length Polymorphism Analysis (T-RFLP).....	31
2.4.3 Quantitative PCR – (qPCR).....	32
2.4.4 Massively parallel sequencing.....	33
2.5 Measuring microbial community diversity.....	34
2.5.1 Diversity and richness estimators.....	36
2.5.2 Shannon’s and Simpson’s Diversity Indices.....	37
2.5.3 Nonparametric Methods to Estimate Diversity.....	38
CHAPTER 3: MATERIAL AND METHODS.....	41

3.1 Site description and sample collection.....	41
3.2 Endophytic bacteria isolations and growth conditions.....	41
3.3 DNA extraction and amplification of 16S ribosomal RNA gene.....	43
3.4 Detection of <i>nifH</i> gene by PCR amplification.....	44
3.5 16S rRNA and <i>nifH</i> sequences analysis.....	45
3.6 Statistical analysis	46
3.7 DNA extraction and Pyrosequencing	49
3.8 Sequencing quality control and Bioinformatic analysis	50
CHAPTER 4: RESULTS.....	52
4.1 Isolation and enumeration of <i>T. esculentum</i> endophytic bacterial population..	52
4.2 Identification and Phylogenetic evaluation of cultured endophytic bacteria associated with <i>T. esculentum</i>.....	57
4.3 Community structures of endophytic bacteria associated with <i>T. esculentum</i> Cluster analysis and diversity indices of Microbial communities.....	62
4.3.1 Cluster analyses	62
4.3.2 Diversity indices	64
4.4 Microbial community analysis of the three sites and plant parts using Principal Component Analysis (PCA).....	68
4.4.1 Organ effect	70
4.5 Phylogenetic analysis.....	77
4.5.1 Phylogenetic analysis of 16S-rRNA sequences.....	77
4.5.2 Phylogenetic analysis of <i>nifH</i> sequences.....	86

4.6 Comparative analysis of bacterial communities associated with different parts of <i>T. esculentum</i> as determined by Pyrosequencing	91
4.6.1 Phylum, class-level analysis of bacterial communities in <i>T. esculentum</i> plant organs.....	91
4.7 Cluster analysis.....	97
4.8 Contributions of Bacterial Taxa to Principal Components.....	99
CHAPER 5: DISCUSSION	113
5.1 Numerical analysis of culturable bacterial communities living in <i>T.esculentum</i>	113
5.2 Structure of culturable <i>T. esculentum</i> endophytic communities.....	116
5.2.1 <i>T. esculentum</i> above ground culturable microbial community diversity.....	119
5. 2.2 Below ground culturable <i>T. esculentum</i> microbial community.....	121
5. 2.3 <i>T. esculentum</i> culturable diazotrophic bacteria	126
5. 3 Microbial community structure.....	130
5.3.1 Cluster analysis.....	130
5.3.2 Principal Component Analysis of microbial communities analysis	131
5.3.4 Phylogenetic analysis	135
5. 4 Bacterial communities associated with leaves, stems and tubers of <i>T. esculentum</i> determined by 454 pyrosequencing techniques	135
5.4.1 Bacterial community structure revealed by 16S r RNA gene amplicon sequencing.	136
5.4.2 <i>T. esculentum</i> -associated bacterial communities.....	137
5.4.3 <i>T. esculentum</i> above ground microbial community diversity.....	139

5.4.4 <i>T. esculentum</i> below ground (tuberous root) endophytic microbiota.....	144
CHAPTER 6: CONCLUSION.....	152
CHAPTER 7: RECOMMENDATIONS	155
REFERENCES	159
APPENDIX : APPENDIX 1: DIVERSITY OF CULTURABLE BACTERIA WITHIN DIFFERENT SAMPLES	193
APPENDIX 2. TABLES CHAPTER FOUR.....	195
Table 12: Sequence similarity between diazotrophic bacterial isolates and the closest type strain from the NCBI Database. A Comparison between the first hit and cultured bacteria.	Error! Bookmark not defined.
Table 12: Sequence similarity between diazotrophic bacterial isolates and the closest type strain from the NCBI Database. A Comparison between the first hit and cultured bacteria (Continued).....	Error! Bookmark not defined.
Table 14. Eigenvalues and percentage of explained variability extracted from <i>T. esculentum</i> sequences based on 16S rRNA gene sequences generated using Roche/454 pyrosequencing.	Error! Bookmark not defined.
APPENDIX 3: Change in diversity of total endophytic microbial communities in leaves, stems and tuber of <i>T.esculentum</i> over time and location. OM, HARN, OTJ represent Omirata, Harnas and Otjinene respectively while the number represent the collection year.....	198
APPENDIX 4. Loadings Matrix values for extracted Principal Components	199
APPENDIX 5: Number and identity of <i>T.esculentum</i> bacterial community identified via the RDP database based on 16S rRNA gene sequences generated using Roche/454 pyrosequencing.....	200
APPENDIX 6: Loading matrix values for extracted principal Component for pyrosequencing.....	204

Appendix 7: Percentage Sequence similarity of selected isolates associated with <i>T. esculentum</i>	208
---	------------

CHAPTER 1: INTRODUCTION

1.1 Background of the study

In almost every ecosystem, microorganisms form the “unseen majority” that exceed by far plants and animals in abundance and diversity (Palmer et al., 2006). Plants offer several ecological niches that harbor diverse communities of symbiotic microbes. They often have beneficial effects on the host, such as improved photosynthetic efficiency, nutrient and water use and tolerance to abiotic and biotic stress (Marasco et al., 2012) and defensive mechanisms (Newton, Fitt, Atkins, Walters, & Daniell, 2010). Some of these microorganisms are best adapted for naturally living inside plants, colonizing inner tissues without conferring pathogenicity to their hosts (Compant, Clément, & Sessitsch, 2010). They are designated as endophytes (Mano & Morisaki, 2008).

The plant endophytic microbial community emerges as an attribute that contributes to the plant’s ability to acclimatize to the environment (Bulgarelli, Schlaeppi, Spaepen, van Themaat, & Schulze-Lefert, 2013). Contrary to the common belief that drylands such as deserts have little potential for agricultural economic activities (Per, Ryden, & Esikuri, 2005), it was recently advocated that significant economic and nutritional advantage can be gained by increasing nutritional use of wild plants (Per et al., 2005) occurring on marginal, arid, and semi-arid lands (Gupta, Panwar, & Jha, 2013). The sustainable exploitation of genetic diversity for improved food security can be achieved with the introduction of new or unexploited so called “orphan crops” (Dakora, et al., 2009).

According to Dakora & Keya (1997) most of these plants' nutritional quality has been assessed and has been found promising.

Tylosema esculentum (Burchell.) A. Schreiber (*Leguminosae*) also locally known as Marama bean is a wild under-utilised and highly nutritious indigenous desert bean. It is found in Namibia (mostly in Omaheke and Otjozondjupa regions), Botswana (around the central Kgalagadi) and to lesser extend in South Africa (North West and Gauteng provinces) (Castro, Silveira, Coutinho, & Figueiredo, 2005). Its seeds are rich in oil and proteins. It forms part of the diet of the indigenous Kalahari people (Holse, Husted, & Hansen, 2010). This plant is also known as the “Green Gold of Africa” (National Research Council, 2006).

T. esculentum has been favored by indigenous people for its nutritional and medicinal value (Chingwaru et al., 2011). This plant has great agricultural potential to produce harvestable material such as seeds and tubers (Holse et al., 2010). It is being considered for domestication (Chimwamurombe et al., 2009). Like other plants thriving in harsh desert conditions, *T. esculentum* might have developed adaptive measures to the desert environment. Still, many fundamental questions about *T. esculentum*'s microbial flora remain unanswered, such as which organisms occupy the diverse specific niches and microenvironments of this plant; how do these bacterial populations vary from individual to individual and over time; how are they affected by geography (Palmer et al., 2006)? Our samples were collected from Omitara and Harnas that are known for their commercial

farming while Otjinene is located in the communal farming area. Some differences in microbial communities were expected.

With the advent of global climate change, understanding *T. esculentum*-microbial community ecological interactions is not only of crucial importance for the plants survival, but also for their biotechnological application (Dakora, et al., 2009). Hence, *T. esculentum*'s microbial, agricultural and biotechnological potential could be of importance for the Southern African region in the search for new resilient crops that can grow in drought-prone, infertile and nitrogen poor soils. Biotechnology has great potential to influence and benefit agriculture. Therefore, integrated and broad microbe management systems that would develop new crop cultivars relying on beneficial functions provided by the plant-microbe symbiosis is a potential alternative to the use of chemical fertilizers in Southern Africa. All this with the goal of increasing agricultural productivity and protecting the environment even on land formerly considered unsuitable for agriculture. *Tylosema esculentum* might be harboring microbes with functional and agronomic potential traits that might out-compete exotic crop species that have taken root in our farming systems (Dakora et al., 2009).

Endophyte studies have mainly focused on plants of economic importance such as rice (Baldani & Baldani, 2005), maize (Montañez, Blanco, Barlocco, Beracochea, & Sicardi, 2012), cotton (McInroy & Kloepper, 1995), Soy bean (Pimentel, Glienke-Blanco, Gabardo, Stuart, & Azevedo, 2006), potato, red clover (Sturz, Christie, & Matheson,

1998) and sugar cane (James & Olivares, 1997) leaving other plants such as marama bean with little to no attention.

If a culture-dependent diversity study can give us a glimpse of the endophytic bacteria diversity, the high-throughput also known as the Next Generation Sequencing Technology will shade light on the bigger picture of the *T. esculentum* endophyte diversity. Further, the metagenomic studies will not only confirm this rich microbiota, but also functions that *T. esculentum* endophytes fulfill in nature. The use of the endophyte would be valuable for sustainable future *T. esculentum* farming and other fragile dry environment prone plants' management.

1.2 Statement of the research problem

T. esculentum, possesses great agricultural potentials to produce harvestable material such as seeds and tubers (Holse et al., 2010). *T. esculentum* seeds have a protein content as high as (30-39%) that is similar to that of soybeans (38%) (Holse et al, 2010), yet this plant presents no apparent nodules. Despite these potential exciting traits, little attention has been given to this plant (Chimwamurombe et al., 2009).

Many research groups around the globe have showed that endophytic bacterial communities have been linked to plant environmental adaptation (Colemann-Dar & Tring, 2014). Thus, it can be predicted that *T. esculentum* harbours a large diversity of endophytic bacterial communities that contribute to its survival in the Kalahari. The composition of *T. esculentum* microbial communities has not been explored. Conversely, *T. esculentum*

bacterial communities' identity, preference colonization (leaves, stems and tubers) and diversity have never been established. Therefore, the knowledge of these niche specific endophytic microbial communities would be essential in clarifying mechanisms involved *T. esculentum* adaptability to harsh conditions.

1.3 Objectives of the study

The present study aims at assessing the total *T. esculentum* endophytes bacterial communities from Omitara, Harnas and Otjinene using culture dependent and culture-independent (Next Generation Pyrosequencing of 16S rRNA gene amplicons) with the goal to analyze microbial richness, diversity and community composition in tubers, stems and leaves. Using culture dependent methods, the diazotrophic bacteria presence occurring in *Tylosema esculentum* will be verified and the *nifH* diversity will be determined, providing a survey of total and diazotrophs endophytic bacteria associated with this plant. The objectives of study were:

- a) To determine and compare tissue specific bacterial density in *T. esculentum*'s leaves, stems and tubers,
- b) To determine and compare tissue specific (leaves, stems and tubers) endophyte bacterial community structure associated with *T. esculentum* (Kalahari Desert);
- c) To determine whether tissue-specific *T. esculentum* associated endophyte bacterial communities are indicative of any nitrogen-fixing functional role within

T. esculentum symbiosis using culture dependent technique;

d) To evaluate the phylogenetic diversity of culturable bacteria endophytes of *T. esculentum* plants growing in the Kalahari Desert;

This study is an indispensable step towards understanding the ecology and interaction between endophytic bacteria and *T. esculentum*.

1.4 Significance of the study

Among prospective climate change-adapted crops for future agriculture, *Tylosema esculentum* (Marama bean), a desert drought avoiding plant, with extraordinary dietary properties, stands out as a prominent contender. Deciphering *T. esculentum*'s putative endophytic microbial communities is of utmost importance. This knowledge is essential for understanding and potentially improving *T. esculentum* – microbe interactions, and further improving its production sustainability. It would be useful in developing adequate means for their biotechnological applications in Marama domestication process. Additionally, *T. esculentum* can make up an informal self-insurance and an alternative source of income to the local communities since it grows in areas where animal husbandry is the sole source of income.

CHAPTER 2: LITERATURE REVIEW

2.1 *Tylosema esculentum* (Marama bean): A plant with great agricultural potential.

2.1.1 Botanical and ecological description of *T. esculentum*

Tylosema esculentum also known as marama bean belongs to the early diverging legume clade, *Detarieae* with species of agronomic interests mostly found in arid regions (James & Olivares, 1997). The genus *Tylosema* (Schweinf.) Torre & Hillc. (*Leguminosae*, *Caesalpinioideae*) has been recently taxonomically evaluated by Castro et al (2005). In their study, it was revealed that the *Tylosema* genus belongs to the *Cercideae* tribe, *Leguminosae* family and the *Caesalpinioideae* subfamily. All its 5 species namely, *Tylosema fassoglense*, *Tylosema humifusum*, *Tylosema argenteum*, *Tylosema angolense* and *Tylosema esculentum* are native to Africa (Figure 1).

Tylosema has a wide range of distribution ranging from Sudan southward to Limpopo, North West and Gauteng Provinces in South Africa and Swaziland (Castro et al., 2005) (Figure 1). *Tylosema esculentum*, is a long-lived perennial, tuberous (Figure 2) (Cannon, May, & Jackson, 2009) and non-nodulating *Leguminosae* (Holse et al., 2010). It grows on well drained poor sandy soils with low water retention capacity. It can also be found on limestone including dolomite soils. The rainfall ranges between 100 mm and

900 mm, and the soil pH between 6 and 8. It is found at an altitude ranging between 1000 m and 1500 m (Mitchell, Keys, Madgwick, Parry, & Lawlor, 2005).

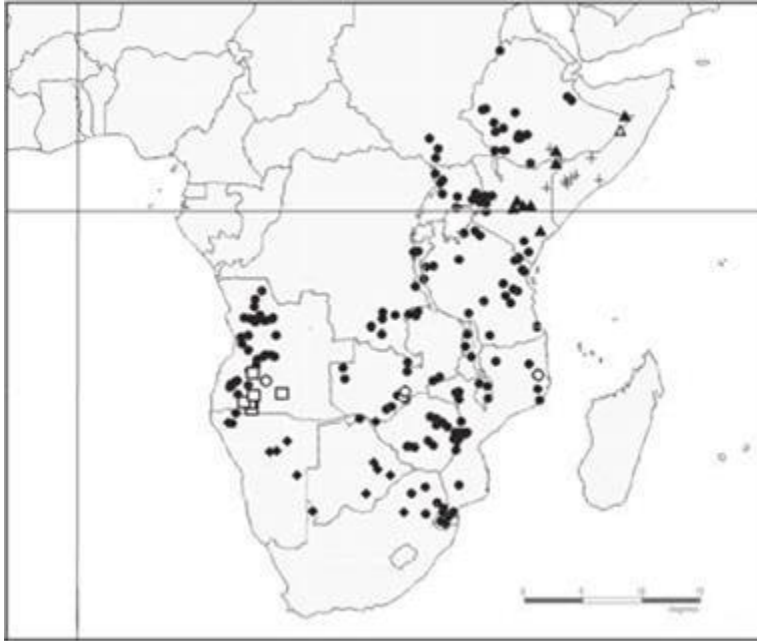


Figure 1: Distribution map of the genus *Tylosema*: *T. angolense* (○), *T. argenteum* (+), *T. esculentum* (◆), *T. fassoglense* (●), *T. humifusum* (Δ, Δ), unidentified specimens (○) adapted from (Castro et al., 2005).

The regional average daily temperatures where *T. esculentum* grows are high with a maximum of 37° C (Mitchell, Keys, Madgwick, Parry, & Lawlor, 2005). *T. esculentum* propagates in flat grassy and wooded grassland in localized patches (Holse et al., 2010).

T. esculentum is a creeper with runners that spread fast in all directions. Under good conditions they can attain 6 m long on the ground (Travlos, Economou, & Karamanos, 2007). They have a green thick mat of leaves that serve to increase light interception (Figure 2) (Travlos et al., 2007). Marama bean's cardian leaflet movements

are partly controlled by the potassium concentration in the soil. This adaptive phenomenon plays an important role on the overall water economy strategy (Travlos, Liakopoulos, Karabourniotis, Fasseas, & Karamanos, 2008) and its leaflet and stomatal movements to save water (Travlos et al., 2007). Equally, these movements are triggered by changes of soil water content, temperature, photon flux density and plant nutritional status (Hartley, Tshamekeng, & Thomas, 2002).



Figure 2: *T. esculentum* (Marama bean) plant with its tuberous root in Harnas (Omaheke region, Eastern Namibia) (A) and *T. esculentum* mature dry seeds (B)

T. esculentum is a drought avoiding plant (Chimwamurombe et al., 2009). When winter sets in, the plant sheds its leaves and stems. It enters a sort of drought dormancy until the next rainy season when it will resprout (Bower, Hertel, Oh, & Storey, 1988).

While under ground, *T.esculentum*'s tuberous root can withstand prolonged long periods of drought. It has developed adaptive strategies whereby it uses its tuber (water content about 85-90%) as a water reservoir (Figure 2). *T. esculentum* uses the short rainy season (mostly three months) to complete its reproductive cycle (Castro et al., 2005)

2.1.2 *T. esculentum*'s nutritional value

T. esculentum is a protein-rich species plant that could be used for boosting food security (Amonsou, Taylor, & Minnaar, 2013). Most studies done on *T. esculentum* have mainly focused on its seeds nutritional qualities (proteins and fat content). It has been established that *T. esculentum* seeds are a good source of proteins (Amonsou, Taylor, Beukes, & Minnaar, 2012).

These studies gave little attention to the fact that *T. esculentum* could be referred to as a functional food (Amonsou, Taylor, & Minnaar, 2013). It is known that diet, until recently, was meant to sustain life and the physical growth only. Nowadays, there is a tendency of taking into account the physiological benefits of the diet that would help individuals to cope or avert diseases (Amonsou, Taylor, Beukes, & Minnaar, 2012). Besides, there is a lack of documentation about the influence of the environment on the vegetative growth as well as the nutritional qualities, the antioxidant activity and the phenolic compounds contained in *T. esculentum* (Amonsou, Taylor, Beukes, & Minnaar, 2012; Amonsou, Taylor, & Minnaar, 2013).

The protein and oil content of *T.esculentum* is comparable to soya and peanuts. Maruatona, Duodu, & Minnaar (2010) confirmed that *T. esculentum* seeds contains 30-39% of protein and 35-43% of oil. Musler & Schonefeld (2006) reported 34.17% of proteins and 39.93% of fat content in *T. esculentum*. These values are in line with previous studies. Amarteifio & Moholo (1998) reported 34.1% of protein and 33.5 of fat content while Mmonatau (2005) reported 30-39% of proteins and 36-43% of fat.

According to Holse et al. (2010), previous studies on Marama bean's nutritional values were based on a small number of samples. They evaluated *T. esculentum* seeds nutritional quality covering the period between 1990 and 2008. These seeds were collected from Botswana, South Africa and Namibia. It was confirmed that their protein and lipid content ranged between 29-38 % and 32-42 % of their dry matter respectively. Their results were in agreement with previous studies (Amarteifo & Moholo, 1998; Hartley, 2002; Maruotona, Duodu, & Minnaar, 2010). The dry matter content (4%) (Bower et al, 1988) and water (3.5%) (Amarteifo & Moholo, 1998) were in accordance with previous studies.

The fatty acids, in *T. esculentum* are unsaturated and 87% are a combination of oleic acid, linoleic acid and palmitic acid (Musler & Schonefeld, 2006). According to Mmonatau, (2005); Bower et al. (1988) and Wang, Vinocur, Altman (2003), in most legume seeds, proteins are the major component as salt globulins or storage proteins. Seed legumes have relatively low sulfur amino acid content, tryptophan and methionine. They are richer in lysine, another essential amino acid (Bower et al., 1988). *T. esculentum* is not

an exception. Compared to legumes, cereals have a greater amount of sulfur amino acid. This makes cereals and legumes nutritionally complementary (Duranti & Gius, 1997).

2.1.3 Uses and potential uses of *T. esculentum*

For many years, the use of *T. esculentum* has remained at a household level and mainly eaten as a snack after being roasted in hot sand (National Research Council, 2006). It is never eaten raw because of its unpleasant taste (Holse et al., 2010). These roasted seeds can be pound or ground to produce a porridge or cocoa like beverage (Holse et al., 2010).

The performance of *T. esculentum* as a probable potential biofuel plant have been recently investigated (Gandure, Ketlogetswe, Temu, 2014). *Tylosema esculentum*'s seeds produced lower concentrations of hydrocarbons in comparison to petroleum diesel (Gandure, Ketlogetswe, Temu, 2014). In addition, *T. esculentum* has high potential as a source of medical and cosmetic products (Chingwaru et al., 2011). Recently, Chingwaru, et al. (2011) showed that *T. esculentum* extracts had high multispecies antibacterial and anticandidal activities comparable to that of conventional drugs. According to the same authors, *T. esculentum* beans and tuber extract were traditionally used against diarrhea in Southern Africa. And, they inferred that these extracts are promising microbicides against Retroviruses.

Mosele et al. (2011) suggested that this underutilized legume is a prospective crop plant with interesting food processing applications. *T. esculentum*'s high protein content

can easily be used as an additive to increase the protein quality of cereal based meals in Southern Africa (Maruatona et al., 2010). It has been shown that the use of *T. esculentum* with Sorghum flour and porridge drastically improved their nutrient quality and antioxidant activity (Kayitesi, De Kock, Minnaar, & Duodu, 2012). However, *T. esculentum* being a bean, its toxicity have not yet been assessed especially for baby formula feeds.

2.2 Endophytic microorganisms

The term endophyte designates all microorganisms that are *senso stricto* at least transiently symptomlessly contained within plant tissues at the time of the study (Reinhold-Hurek & Hurek, 2011). Because of the presence of photo-assimilated compounds, endophytes establish a mutualistic relationship with plants to gain extended nutrient and defensive mechanisms (Gomes & Cleary, 2013).

Many researchers Rosenblueth & Martínez-Romero (2006); Hardoim, van Overbeek, & Elsas (2008); Compant et al. (2010); and Reinhold-Hurek & Hurek (2011) suggested that not only endophytes should be isolated from surface sterilized plant tissues but they also mentioned the endophyte status should be confirmed through the Koch's postulates. Therefore, they have to be microscopically visualized within the plant tissues. The definition, though widely accepted, needs to be upgraded for molecular advances have shown that some endophyte microorganisms can only be sequenced not isolated. Rout &

Southworth (2013) suggested a more inclusive definition of endophytes. They define endophytes as sets of microbial genomes found inside plant organs.

2.2.1 Mechanisms stabilizing plant-microbe symbiosis

Two mechanisms are responsible for symbiont transmission, namely vertical transmission or horizontal transmission. However, some hosts employ both modes of transmission (Rosenberg & Zilber-Rosenberg, 2011). Microbial symbiotic associations exhibit a non-random partner association (Berlec, 2012). Hosts display control mechanisms to select against ineffective bacteria symbionts (Partida-Martínez & Heil, 2011) to align partners' interests and sanction against cheaters (Angus & Hirsch, 2015).

2.2.2 Vertical or transovarian transmission

According to Bordenstein & Reznikoff (2005), vertical symbiotic bacteria are inherited from female to offspring at the time of host reproduction through the maternal eggs. In this case, partners are intimately interdependent and cannot be separated (Lund, Kjeldsen, & Schramm, 2014). They establish an obligate and irreversible association with their hosts. The endosymbionts improve host fitness (Bordenstein & Reznikoff, 2005) which can go to the extent of substituting its own immune function before the comprehensive development of the host's immune system (Pradeu, 2011).

Transovarian transmitted symbionts are of low diversity, effective and pressured to maintain and enhance their effectiveness because they are not only obligate but depend on their host fecundity (Douglas, 1998). They can easily be experimentally manipulated within the host population without contagious spread (Yule, Miller, & Rudgers, 2013).

It has been suggested that vertical endosymbiont transmission is a mechanism towards a stable evolution of a mutualistic relationship (Bordenstein & Reznikoff, 2005; Verstraete, Janssens, Lemaire, Smets, & Dessein, 2013). This type of transmission has been found in a wide range of microbes involved in beneficial effects to their host (Verstraete et al., 2013). Vertically transmitted symbionts are better spread by healthy hosts for their presence should lead to the evolution of strategies of avoidance or sabotage with virulent horizontally transmitted parasites (Vautrin, 2009).

2.2.3 Horizontal or environmental microbial transmission

Horizontally transmitted symbionts are taken up from the environment anew by each host generation (Bright & Bulgheresi, 2010) or contagiously from other hosts in the same generation (Drown, Zee, Brandvain, & Wade, 2013). Hosts are exposed to multiple symbiont genotypes with different degrees of effectiveness (Douglas, 1998). Horizontal transmission favours traits that maximize the fitness of one of the partners without regard to the fitness of the other (Drown, Zee, Brandvain, & Wade, 2013). These symbionts usually do not necessarily live inside the host and are not strictly necessary to the host survival (Drogué, Doré, Borland, Wisniewski-Dyé, & Prigent-Combaret, 2012). They are

capable of engineering their own invasion in the host cells and tissues (Dale & Moran, 2006).

2.3 Plant microbial interaction and resource partitioning

2.3.1 Plant – microbe interaction

As primary producers in most terrestrial ecosystems, plants play important roles as ecosystem engineers (Vacheron et al., 2013). They assemble a specific subset of microorganisms into plant associated microbial communities both in above ground and below ground organs (Lebeis, 2014). To sustain all forms of life including bacteria, plants use energy from sunlight to synthesize various organic compounds that serve as substrate for energy production (Hardoim et al., 2008). In addition, Bisseling, Dangl, and Schulze-Lefert (2009) stated that the plant provides an estimated 21% of its net photosynthetic products to sustain its associated microbial community. Besides gaining extended nutrient access they also benefit from an extended defensive mechanism (Kawaguchi & Minamisawa, 2010).

In semi-arid areas such as Namibia, plant microbe association is important, a necessity and an opportunity for both plants and microorganisms' mutual benefit to adapt to this unpredictable harsh environment (Hartmann, Schmid, Tuinen, & Berg, 2009). It has been documented that symbiosis between bacteria and eukaryotes has played a major role in shaping their evolution and diversity (Sessitsch et al., 2012; Marasco et al., 2012).

They argue that despite the importance of root associated microbiota in stress management in semi-arid environments, studies have not yet received enough attention especially the theory that desert farming promotes and selects microbes that improve key primary services in drought environments.

All plants sense and transmit signals and respond to abiotic stresses. Why are so few of them colonizing high stress habitats (Redman et al., 2011)? Redman and his team (2011) suggested the presence of habitat adapted endophyte microbiota increases the plant fitness to abiotic stress and even form a strategy to mitigate the impact of climate change. Lobato, Silveira, Costa, and Neto (2013) proposed that plant exudates constitute one of the factors that controls the endophyte microbiota population in times of stress.

Rodrigues et al. (2008) studied native grass species from coastal and geothermal habitats. They concluded that symbiotically conferred stress tolerance is a habitat specific phenomenon. They later defined habitat-specific, symbiotically-conferred stress tolerance as habitat-adapted symbiosis. Later, it was hypothesized that this phenomenon is responsible for establishing plants in high stress habitats.

Most plant phenotypes in nature are the result of the combined and coordinated expression of both plant and microbial genes. In fact, much parts of the plant phenotype epitomize the “extended phenotype” of one or several microorganisms (Partida-Martínez & Heil, 2011). These interactions establish a highly effective phenomenon of adaptation that shape microbial ecological communities (Russo et al., 2012).

2.3.2 Microbial niche portioning: a response to resource availability

Several recent studies confirm that plants offer several ecological niches that harbor diverse communities. They often have worthwhile effects on the host, such as strengthened photosynthetic efficiency, nutrient and water use and tolerance to abiotic and biotic stress (Rey & Schornack, 2013). These beneficial traits are comparable to ones searched for by plant scientists while working to develop ecologically sustainable crops for food, fiber and biofuels (Compant et al., 2010). Instead of genetically engineered plants that have gained environmental resistant stress traits attributes, Plant Growth Promoting Bacteria (PGPB) may be used to perform the same roles as suggested by Glick, (2012). However, Ernst, Mendgen, & Wirsal (2003) claim that factors shaping the endophyte microbial diversity distribution are not yet well explained. Two hypotheses trying to establish this causality exist.

Traditionally, it is believed that niche differentiation is the result of biotic and/ or abiotic factors. Concerning microbial endophytes such factors are not easy to establish. Thus, the neutral hypothesis suggests that niche differentiation is the result of irregular factors. Their respective interactions mediate cycles that are valuable to both partners such as nutrient uptake so essential to their health and development and protect hosts against pathogens and herbivores (Tsiamis, Karpouzias, Cherif, & Mavrommatis, 2014).

The bacterial population coexistence can be mediated through nutritional resource partitioning (Newton, Gravouil, & Fountaine, 2010). Niches differentiation for resource

acquisition allows the coexistence of a high diversity of microbes in the same niches. They are probably adapted to use the same resources but at different concentration to avoid damaging interspecific competition (Mayali, Weber, Mabery, & Pett-Ridge, 2014). Several examples of resources partitioning have been observed in cases whereby potential competitors divide resources in space, time or morphologically (Wehner, 2013). Kernaghan (2013) suggested that resource partitioning can also be context-dependent where a given resource is made available to microbes depending on environmental conditions.

According to the stochastic hypothesis, microbes in a given niche are competitively equivalent; thus resources are equally shared among microbes (Wehner, 2013). However, it can be argued that microbial co-colonization of the same niche might be favored by the presence of multiple sources of carbon (sugars) realized by the host plant. This makes it possible for plants to translocate their wide variety of photosynthates to the different organs where they are needed. A specific plant can selectively benefit from providing resources where they are only needed to better control their symbionts (Wehner, 2013).

2.3.3 Plant microbial communities

The bacterial diversity is not a static concept (Andreote, Azevedo, & Araújo, 2009). In a given ecosystem, various ecological biotic and abiotic factors shape the structure and

dynamics of terrestrial plant communities. Afterwards, among biological forces that structure a plant community, plant microbe- interactions involving costs and benefits to each partner, until recently, has received little attention (Knelman et al., 2012).

Equally, the ecological mechanisms governing the association and the structure of these intricate microbial communities and their relationships to the plant and soil community are poorly understood (Knelman et al., 2012). The nearly 300 000 plants species that exist on earth constitute complex and diverse spatially and temporally rich niches for microbial communities (Raaijmakers, Paulitz, Steinberg, Alabouvette, & Moëgne-Loccoz, 2009). They are harbored inside (as endophytes) or on external surfaces (as epiphytes) of their various organs. They make a remarkable diversity of microbial communities known as plant microbiota (Rout & Southworth, 2013). All plant taxa surveyed so far are not an exception to the rule. They are all associated with well-established microbial communities (Compant, Duffy, Nowak, Cle, & Barka, 2005), suggesting that plant-microbe interactions are of ubiquitous nature (Partida-Martínez & Heil, 2011).

According to Lopez, Tinoco-Ojanguren, Bacilio, Mendoza, & Bashan (2012), plant associated microorganisms can be classified according to their effect on plants and the way they interact with host plants. The interaction can be pathogenic, neutral or commensal. The extent to which the host and /or tissues determine the endophyte structural and functional diversity community have not been fully explained (Sun et al., 2008).

Patterns point towards some abiotic factors such as the soil types combined with its edaphic variables and climate and seasons. They are assumed to explicitly influence plant microbial communities (Compant et al., 2010). Conversely, according to Casas, Omacini, Montecchia, & Correa (2011), biotic interactions between living plants and above ground organisms can influence below ground communities with important consequences on the ecosystem functioning.

This triumvirate constitutes the driving force in the plant-microbe interaction (Pedraza et al., 2009). The mycelia exudate can also shape the root bacteria assemblage as it was demonstrated by Artursson, Finlay, & Jansson (2006) when, in a pot experiment, he added arbuscular mycorrhiza (AM) fungi *Glomus mossseae* in soil. At a community level, the strong influence of AM has been reported (Singh, Pandey, Kumar, & Palni, 2010). Similarly, microbes have tremendous influence on the plant's community diversity and productivity (Morrissey, Dow, Mark, & O'Gara, 2004).

Only a small minority of soil microbes will have the privilege to enter the root system of their host to have a more specific enhanced beneficial effect with an endophytic lifestyle as symbionts (Hirsch, Mauchline, & Clark, 2010). Microorganisms engage themselves in more than one mutualism at once. Plants are capable of getting into mutualistic relationships with mycorrhizal fungi, endophytic fungi and beneficial bacteria (Hirsch, Mauchline, & Clark, 2010).

Researchers tend to focus on pair-wise mutualisms neglecting the complex interaction that naturally occur within the plant (Keenan, Rudgers, & Jennifer, 2008).

Plant-microbe interaction ranges from two partite symbioses (legume-rhizobia) to multipartite endophyte, rhizosphere and phyllosphere communities. Research has revealed great taxonomic and functional diversity (Tikhonovich & Provorov, 2011). The most common taxa of isolated heterotrophic endophytes include *Bacillus* (Bai, Zhou, & Smith, 2003), *Enterobacter* (Benedito et al., 2008), *Pseudomonas* (Pereira, Ibáñez, Rosenblueth, Etcheverry, & Martínez-Romero, 2011), *Serratia* (Berg et al., 2002), and *Streptomyces* (Sessitsch, Reiter, & Berg, 2004).

2.3.4 Plant endophytic microbial communities: Assemblage and diversity

Microbes play an integral and often unique role whether directly or indirectly in processes as diverse as biosphere history, functioning and preserving its productivity and providing ecological services to human (Zhong et al., 2011 & Yu-Qing et al., 2008). Contrary to the common belief, according to Tsiamis, Nikolaki, & George (2013), microbes are the major contributors to our planet's photosynthetic capacity, compared to plants. They argue that microbial activities in our ecosystems are not executed by individual microbes but in tune with microbial communities that are endowed with the capacity to acclimatize and maintain themselves even under life-threatening or inhospitable environments (Zhou et al., 2013). Their diversity and community structures remain poorly understood (Sun et al., 2008). They include the number and composition of the genotype, species functional types and units in a given system (Barrios, 2007).

Host associated endophyte communities have been found to be essential for plant health and other relevant ecological roles to the host (Reinhold-Hurek & Hurek, 2011). Using modern sequencing approaches, mechanisms by which microbes exert these effects are being clarified (Berg, Grube, Schlöter, & Smalla, 2014). Deciphering the principles underlying endophyte microbial communities compels detailed analyses using the concept of diversity or that of community structures (Robinson, Bohannan, & Young, 2010).

Throughout their lifetime, plants in their natural habitats are associated with an extremely complex microbial diversity (Dudeja, Giri, Saini, Suneja-Madan, & Kothe, 2012) that includes mycorrhizal fungi, bacterial and fungal endophytes (Craig et al., 2011). This association is referred to as plants' second genome (Bulgarelli et al., 2013) or microbiome (Borer, Kinkel, May, & Seabloom, 2013). Thus, plants and prokaryote associations led to a paradigm shift view of plants as meta-organisms or holobionts rather than an isolated entity (Berg, Grube, et al., 2014) or individuals (Fierer et al., 2012). The plant microbiome can easily be divided in two parts: (a) the below ground where the plant especially tubers, roots and rhizomes are in permanent contact with the soils; (b) the plant part above ground also known as phyllosphere comprises of flower, buds, stem and leaves

The biological significance of symbiosis has been summarized in three respects: (a) it is considered a source of novel capabilities, (b) it plays a determinant role in the eukaryotic evolution and (c) it is a source of the ecological success of some plants and animals (Shelef et al., 2013). These discoveries have deeply challenged the accepted idea

of autonomous individuals (Scott, Sapp, & Tauber, 2012) and its corollary internalist perspective theory (Pradeu, 2011).

As a result, in nature, plants would develop mechanisms that favor effective endosymbiont acquisition (Schenk, Carvalhais, & Kazan, 2012) to exploit future cooperative fidelity and partnerships (Kozyrovska, 2013). Plant-endosymbiont associations complement metabolic means and services to supplement hosts in poor environments to improve their fitness (Reisberg, Hildebrandt, Riederer, & Hentschel, 2013).

Symbiosis between plants and microbes can affect the whole plant health for it can involve complex microbial community interactions (White et al., 2012). Endosymbiont microorganisms are essential for nutrient acquisition and their transport. *De facto*, they would influence crop yield and quality. They are endowed with the capabilities to suppress diseases and promote resistance to different stresses. More importantly, they occupy space that would also be open to pathogens (Borer et al., 2013). Partners' behaviours vary whereby symbionts behaviour can vary from mutualistic to parasitic and the host can "cheat" their symbionts through enslavement (Kozyrovska, 2013). Some genera are ubiquitous and distributed over the entire plant such as *Bacillus* and *Pseudomonas* (Berg, Krechel, Ditz, Sikora, Ulrich, 2005). Some are specifically distributed in different plant microhabitats (Köberl, Schmidt, Ramadan, Bauer, & Berg, 2013). In turn, plants favour a certain microbiome by providing an energy source in the form of carbon-rich

rhizodeposit, adjusting the soil pH and reducing competition for beneficial microbes (Santhanam, Groten, Meldau, & Baldwin, 2014).

2.3.5 Plant Microbial communities fine-tuning

According to Saunders, Glenn, & Kohn (2010), community ecology theory stipulates that ‘filters’ mediate community assembly through a series of processes that result in the coexistence of particular species at a site. It was shown that these processes can roughly be divided into two categories: habitat filtering, and species interactions. Root associated bacterial assemblies occur at the rhizosphere level. They are defined largely in response to exchange of signals triggered from the plant-microbe interaction. It is done in combination with the soil type and the host genotype responsible for the fine tuning of community structure during the establishment of the root endophyte microbiota (Bulgarelli et al., 2013; Lebeis, 2014). This capacity to colonize roots is a major step for successful plant endophyte colonization. According to Compant et al (2005) plant endophyte colonization involves bacteria recognition, adherence, successful colonization and growth. Motility coupled with chemo-attractants play an important role in endophytic plant colonization (Lebeis, 2014).

Bulgarelli et al (2013) extensively worked on bacteria communities in *Arabidopsis thaliana* endophytic and rhizosphere microbiomes using 454 sequencing (Roche) of the 16S rRNA gene amplicons. Plants were grown in pots under controlled gnotobiotic

conditions. In both studies comparison of species richness was done. It was revealed that the soil type defined the composition of the rizosphere microbiome, and the endophytes were a smaller, distinct group (50% fewer species identified than in the rhizosphere) dominated by *Actinobacteria* followed by *Proteobacteria*; *Firmicutes*, *Bacteroidetes* and *Cyanobacteria*.

According to Reinhold-Hurek, & Hurek (2011), *Archaea* do not associate with roots, however they are found on older root surfaces but have not convincingly been found in internal tissues. Yet, Hampp, Hartmann, & Nehls (2012), found that there is abundant evidence that *Streptomyces* colonize plant root surface and even plants tissues and they might be responsible for antibiotic production for host plant protection against phytopathogens. They even emphasized that recently some pathogenic *Streptomyces* species have been described.

2.3.6 Benefits of microbial community diversity analysis

“The role of the infinitely small in nature is infinitely great” opinioned Louis Pasteur. Indeed, all biological cycles and systems in the biosphere depend on microbial activity (Jennifer et al., 2004). Therefore the utmost importance of defining microbial community interactions and their consequences arise (Konopka, 2009). They govern ecological functions, for example, carbon cycling, greenhouse gas emission and oxygen production (Foster, Bunge, Gilbert, & Moore, 2012). These interactions are responsible for essential

biological processes in plant development and health status (Andreote et al., 2009). Remarkably, these activities are not performed by individual microbes but by organized microbial communities that can adapt and excel even under hostile environments (Tsiamis et al., 2014). In plants, endophytes form integrative indispensable partners that are inclined to evolve with time and space (Schulz & Boyle, 2006).

The capacity of an ecosystem to resist extreme perturbations or stress conditions, can partly be dependent on the diversity within the system. According to Fakruddin and Mannan (2013), diversity analyses are therefore important to improve the knowledge of genetic, functional distribution of microbial communities and resources over time (Fakruddin & Mannan, 2013). Additionally, microbial community profiling would allow a thorough understanding of the establishment of patterns associated with health and diseases (Siqueira, Sakamoto, & Rosado, 2010). Similarly, the ecosystem functioning and sustainability depends on preserving a specific diversity (Fakruddin & Mannan, 2013).

2.4 Plant Microbial Diversity Analysis

Microbial biomass and diversity assessment can be estimated by direct microscopic counts and other biochemical counts such as plate counting, phospholipid analysis, chloroform fumigation, ATP enzymatic and enzymatic activity (Andrew et al., 2012). These methods provide an estimate for the active population but they can not provide information on the total population level (Muyzer, 1999; Andrew et al., 2012).

According to Fischer, Fischer, Magris, & Mori (2007) for long, microbial diversity was grossly underestimated using culture dependent technique under standard laboratory conditions. These methods are met with strong limitations (Hardoim et al., 2008). Some bacteria can not grow because culture conditions are not met. Others depend on other microbes for growth. Even so, there is no guarantee the activity measured in the laboratory is relevant in the environment (Tsiamis et al., 2014).

However, stressed or weakened cells need to recover under specific culture conditions to growth in normal conditions (Justé, Thomma, & Ievens, 2008). As a result, the growth will only represent 1% of all microbial species (Rastogi & Sani, 2011). This number can rise to 10% in high nutrient availability conditions such as the rhizosphere (Hirsch & Mauchline, 2012). These limitations prompted the introduction of culture-independent methods (Justé et al., 2008).

The applications of DNA/RNA based techniques have been favored for microbial diversity studies in natural habitats (Barriuso, Valverde, & Mellado, 2011). These techniques have a great impact on the study of functional biological investigations (Delseny, Han, & Hsing, 2010). However, the advances in culture-independent nucleic acid-based methods to characterize microbial communities are fundamentally being unseated by sequence-based analyses of microbial diversity (Fierer & Lennon, 2011).

In amplicon sequencing, a particular gene, gene fragment or sequence is identified and amplified from complex environmental samples. The sequences will be determined (Di Bella, Bao, Gloor, Burton, & Reid, 2013) to allow the analysis of bacterial diversity

satisfactorily (Barriuso et al., 2011). The SSU rRNA gene is present in all living organisms (Palmer et al., 2006). It contains different species-specific domains. It is a widely and accepted conventional tool for identifying and quantifying microbial community composition (Palmer et al., 2006). The 16S rRNA gene (rDNA) amplicon is used as a standard approach to infer phylogenetic affiliations at various taxonomic levels in numerous habitats (Di Bella et al., 2013). Since its use, it has been shedding light on previously unimagined diversity of microbes with unprecedented resolution (Sun et al., 2008).

The bacterial community composition is estimated by clustering polymerase chain reaction of the conserved regions flanked by hyper-variable regions of the 16S rRNA (SSU rRNA) (Yu-Qing et al., 2008). Characterizing the microbial composition enables a microbial community profiling and contributes to understanding the potential impact of a particular microbial community into the ecosystem function and stability (Armougom & Raoult, 2009; Peterson, Frank, Pace, & Gordon, 2010). In natural environments, microbial communities are complex and their diversity has been shown to be difficult to assess and to compare. Hence, several analytical fingerprinting techniques have been developed to allow a rather simple comparison of microbial communities between samples (Parks & Beiko, 2012).

These techniques include denaturing/temperature gradient gel electrophoresis (DGGE/TGGE), terminal restriction fragment length polymorphism (T-RFLP), and quantitative polymerase chain reaction (qPCR) (Su, Lei, Duan, Zhang, & Yang, 2012).

They produce community fingerprints based on either sequence polymorphisms or length polymorphism (Muyzer, 1999). These fingerprint techniques allow a simultaneous multiple sample analysis for difference, similarity or diversity identification. However, they can not identify individual members of the microbial communities (Valášková & Baldrian, 2009).

2.4.1 Denaturing Gradient Gel Electrophoresis (DGGE) and Temperature Gradient Gel Electrophoresis (TGGE)

DGGE or TGGE rely on separating double stranded 16S rRNA gene fragments amplified from environmental DNA samples. The separation principle for both methods is based on the sequence variation within the DNA fragment. It creates differences in mobility of the partially melting DNA molecule on linear gradients of DNA denaturant agents (mixture of formamide and urea in DGGE) or temperature (rather than chemical denaturants) in TGGE on polyacrylamide gels (Muyzer, 1999; Valášková & Baldrian, 2009; Větrovský & Baldrian, 2013). Both DGGE and TGGE involve a special GC rich sequence attached at 5' of the forward primers during the PCR step. This is essential to prevent the two DNA strands from complete dissociation into single strands during electrophoresis (Větrovský & Baldrian, 2013). Gradient gels can be used to identify species in a community through excision and sequencing individual bands of interest (Štursa et al., 2009).

The ability to only resolve most dominant or top taxa in a community constitute one of the major drawbacks. Similarly, band from more than one species may be hidden behind a single band resulting in underestimation of bacterial diversity (Muyzer 1999). DGGE/TGGE is especially useful when a microbial community is dominated by a few members, as time points or multiple samples can be represented in one gel image and easily visually compared through presence/absence or intensity of bands.

2.4.2 Terminal Restriction Fragment Length Polymorphism Analysis (T-RFLP).

Terminal Restriction Fragment Length Polymorphism analysis has been used to monitor changes in the structure and composition of microbial communities (Liesack & Dunfield, 2004). It has been applied to the composition and diversity analysis of soil (Kitts, 2001) and marine and activated sludge system microbial communities (Eschenhagen, Schuppler, & Röske, 2003).

A combination of digested fluorescent end-labelled PCR products and the automated sequence technology form the pillar of this technique (Hirsch et al., 2010). This fingerprinting technique uses either the 5' PCR primer, or both primers, labeled with a fluorescent marker. The resulting set of T-RFLP of a community is a number of T-RFs known as T-RF profile (Fredriksson, Hermansson, & Wilén, 2014). The restriction fragments (T-RFs) are separated by polyacrylamide or capillary gel electrophoresis with laser detection of the only labelled fragments using an automated sequencer that provide

suitable digital output (Fredriksson et al., 2014). They are viewed as an electrogramme showing the microbial community profile as a series of peaks of varying height (Tiquia, 2010). Despite its many benefits and useful applications, RFLP analysis is still a slow and more tedious compared to some of the newer DNA analysis techniques. In addition, it also requires substantially larger sample sizes than other forms of analysis

2.4.3 Quantitative PCR – (qPCR).

Quantitative real-time PCR (q-PCR) is another molecular technique that can be used to study microbial communities. Differing from traditional PCR, which relies on end-point detection of amplified genes, q-PCR is fluorescence based. It is a highly sensitive tool used for gene expression quantification and expression profiling. By incorporating an intercalating fluorescent dye such as SYBR Green or fluorescent probes (TaqMan); DNA quantification is monitored in real time (Andreote, Azevedo, & Araújo, 2009). Therefore, data is collected through the PCR process as it occurs, thus combining amplification and detection into a single step. Q-PCR is highly sensitive to starting template concentration (Rastogi & Sani, 2011).

It has, therefore, the advantage to detect a limited amount of starting material (Gachon, Mingam, & Charrier, 2004) and it allows to determine the amount of starting template concentration before the amplification (Wong & Medrano, 2005). One of the main advantages of real-time PCR is its rapidity to provide reliable data. However, this

technique is limited by the number of samples that can be directly tested as specific probes are designed for each bacteria of interest. In comparison to gene cloning and sequencing which are labor-intensive, time-consuming and expensive, molecular fingerprinting methods show more advantages such as rapidity and reproducibility and a bigger number of samples can be analysed.

With these techniques, huge steps were made in analyzing the complex microbial community structure, composition, spatial distribution and dynamics in the ecosystem (Andreote et al., 2009). With the advent of the promising high throughput massively parallel sequencing technologies, enough sequencing depth to cover the complex microbial communities were achieved.

2.4.4 Massively parallel sequencing

The much faster Next Generation Sequencing (NGS) also known as High-Throughput Sequencing (HTS) is cost-effective. It avoids the laborious traditional culture-dependent and DNA fragment cloning studies (Ansorge, 2009). This novel sequencing technology has revolutionized our ability to study complex environmental microbial communities' compositional features. They include gene content, functional significance and diversity either by shotgun metagenomic or amplicon sequencing approaches (Tsiamis et al., 2013). The NGS technique generates massive amounts of data in a single run which leads to a rapid in-depth analysis (Inceoğlu, Al-Soud, Salles, Semenov, & van Elsas, 2011).

The NGS amplicon data can be performed using metabarcoded 16S rRNA taxonomy-dependent or independent algorithms to identify different species and compare different communities (Dumont, Lüke, Deng, & Frenzel, 2014). The taxonomy-dependent methods rely on annotated sequences existing in reference databases (Cai & Sun, 2011). BLAST is the most commonly used algorithms when classifying sequences (Cai & Sun, 2011). Several well-curated public databases such as Ribosomal Database Project (RDP), Greengenes (Wang, Cai, Sun, Knight, & Mai, 2012), SILVA (Quast et al., 2013) exist.

Reads with less than a predetermined threshold are assigned to a matching reference sequence in a domain, phylum, class, order, family, genus and species (Dumont et al., 2014). In this approach, a serious drawback exists. Most of microbes are still unknown or do not have their 16S sequences included in the database. It makes it difficult to classify novel sequences (Chen, Zhang, Cheng, Zhang, & Zhao, 2013). In addition, they claim that in these reference databases, most microbes are only well characterized at the genus level. To identify individual microbe sequences is of utmost importance in microbial ecology. The basic information such as its traits, physiology, epidemiology and evolution history can help to clarify their ecological role in a given environment (Kim et al., 2013).

2.5 Measuring microbial community diversity

Diversity is a general ecological concept (Bent & Forney, 2008). Its persistence and composition in an ecosystem depend on the interaction of resource availability and biotic

and abiotic factors (Jost, 2010). The widespread use of culture independent high-throughput sequencing (HTS) technologies coupled with bioinformatics tools has made it possible to provide powerful means of surveying complex microbial communities and quantifying their microbial diversity (Morgan & Huttenhower, 2012). Though challenging, microbial community diversity measures are seen as an important indicator of microbial ecological system stability (Zhang et al., 2012).

Such understanding in microbial ecology is a challenge for there are no universally accepted discrete boundaries between functional genes, between species genes and even the species concept is still under debate (Caro-Quintero & Konstantinidis, 2012). Currently, a particular interest is being devoted to understanding which factors shape microbial community diversity patterns and afterwards the reliability of estimators used (Pacchionni et al., 2014). Most microbes are yet to be cultured. Nonetheless, the widespread metagenomic studies based on High Throughput DNA Sequencing (HTS) has created a sequence data deluge that has revealed extraordinary diverse species in natural environments (Li, Bihan, Shibu, & Methé, 2012).

High bacterial diversity has been described in several habitats. These habitats are as diverse as those of human skin (Fierer & Lennon, 2011; Morgan & Huttenhower, 2012). Bodenhausen, Horton, & Bergelson (2013) studied *Arabidopsis thaliana*'s root and leaves. Köberl et al. (2013) and Fierer & Lennon (2011) worked on plant of medicinal and economic importance and the polar desert in Antarctica respectively. Several phylogenetic markers have been used to characterize microbial communities and statistical indices have

been applied to quantify microbial diversity and test hypothesis regarding patterns and processes in microbial communities (Bent & Forney, 2008).

2.5.1 Diversity and richness estimators

Species diversity within a particular area, habitat or sample is often referred to as alpha (α) diversity (Di Bella et al., 2013). It provides an assessment of the total microbial community richness in a single sample or environment (Armougom & Raoult, 2009). When analyzing microbiological sequences, out of necessity, raw sequences are processed and assigned to artificial proxy species called Operational Taxonomic Units (OTUs) (Bohannan & Hughesy, 2003). Nowadays most alpha diversity are being used as statistical examination of microbial samples collected from the environment (Bohannan & Hughesy, 2003).

The term diversity consists of two properties that are considered when measuring diversity (Bohannan & Hughes, 2003):

(a) The richness (S) or abundance is defined as the number of different OTU types present in a given community, habitat or sample. In other words, it refers to the quantitative variation among OTU types (Hughes & Bohannan, 2004). This is the simplest measure of community diversity for it does not involve any information about the relative abundances of different OTU types (Fakruddin & Mannan, 2013).

(b) In contrast, the second component, evenness or equitability, describes the relative individual distribution or abundance or biomass among OTU types (Hughes & Bohannan, 2004; Zhang, et al., 2012).

These definitions are easily applicable to macrobiotics (Koskinen, Hultman, Paulin, Auvinen, & Kankaanpää, 2011). One problem is that evenness often is unknown in bacterial systems because individual cells are rarely identified to the species level (Fakruddin & Mannan, 2013). According to Fakruddin & Mannan, (2013), microbial diversity should be considered at three levels: within species (genetic), species number (species) and community (ecological) diversity. Communities with OTUs that are more divergent from each other are considered more diverse (Lozupone & Knight, 2008). To estimate and compare species biodiversity, several statistical indices exist. The most popular are Shannon index and Simpson index.

2.5.2 Shannon's and Simpson's Diversity Indices

The Shannon's diversity index is the most commonly used index of heterogeneity. It measures the average degree of uncertainty in predicting as to what species an individual chosen at random from a collection of S species and N individuals will belong. The value increases as the number of species increases and as the distribution of individuals among the species becomes even. The Shannon's diversity index is the most commonly used index of heterogeneity to calculate the average degree of uncertainty inherent in

abundance distribution. At the same time, it predicts which OTU would be if that individual were picked at random from the community (ie as each OTU in a community comes closer to having the same number of individuals) (Eppelleberg & Fedor, 2003). It is also known as Shannon-Wiener index, the Shannon-Weaver index and the Shannon entropy. The Shannon index varies from 0 for communities with one species to various other values for other species mixes. The Shannon –Weiner Index is used to compare diversity between different samples or environments.

It is positively correlated with both species richness and evenness of OTUs (Hill, Walsh, Harris, & Moffet, 2003) and it indicates the uniformity of species and its abundance in OTUs (Keshri, Mishra, & Jha, 2013). The Shannon's diversity index though popularly used, presents some drawbacks (Hill, Kerry, Walsh, Harris, & Moffet, 2003). It is not a probabilistic measure of the difficulty in predicting the identity of the next bacterial clone, but it points out how a particular sample has the highest H' subsequently is the most diverse (Hill et al., 2003). It provides more information about community composition, placing a greater weight on species richness (Schoss & Handelsman, 2004).

2.5.3 Nonparametric Methods to Estimate Diversity

Nonparametric diversity estimators use the incidence or abundance of rare species in the sample to estimate the total number of species (Basualdo, 2011). There is no exhaustive sampling method that would capture the total microbial community richness (Bent &

Forney, 2008). The richness estimation is always based on available biological inventories (Basualdo, 2011). The maximum diversity richness is determined by the number of OTUs present which can be influenced by the sequencing depth. Therefore, the observed richness would always be lower than the true richness (Koskinen et al., 2011). To estimate the expected maximum richness, qualitative based richness estimators such as rarefaction curves, Chao1 (Chao, 1984), ACE (abundance-based Coverage Estimator) (Chazdon, Colwell, Denslow, & Guariguata, 1998) and Jackknife are applied (Koskinen et al., 2011). Nonparametric estimators are particularly useful for small sample size because they make no assumption about the underlying distribution (Curtis et al., 2006).

Chao1 estimator is a richness estimator and it relies on the presence of singletons (OTUs represented by only an individual in a sample) and doubletons (OTUs represented by two individuals in a sample) (Bohannan & Hughes, 2003). This estimator gives an estimate of the minimum diversity compatible with the data (Curtis, et al., 2006). If there are no singletons and doubletons in the sample, Chao1 equals the number of observed species (Curtis et al., 2006).

The suitability of this estimator is still under scrutiny especially for pyrosequencing of 16S rRNA, since the viability of singleton and doubletons as representative of a true species in the sample is still under debate. With the exception of singletons and doubletons, the Abundance-based Coverage Estimator (ACE) considers abundant species that are represented by more or fewer than 10 individuals (Basualdo, 2011).

Rarefaction is a technique to assess the species richness expected from the number of sequences and defined OTUs (Keshri, Mishra, & Jha, 2013). It compares observed taxon richness among sites, treatments or habitats that have been unequally sampled or have not been fully sampled (Hughes & Bohannan, 2004). It is accepted that large samples have more OTUs than small samples even though they are drawn from the same community (Hughes & Bohannan, 2004).

CHAPTER 3: MATERIAL AND METHODS

3.1 Site description and sample collection

The study sites were located in Omitara (S22° 21.596' E18° 02.476') in Khomas region, Harnas (S21° 47.705' E19° 19.921') and Otjinene (S21.1333°, E18.7667') in Omaheke region in the Eastern part of Namibia. Omaheke is situated on the eastern border of Namibia with Botswana and forms the Western extension of the Kalahari Desert.

Once a year from 2011 to 2014, during the rainy season, samples were randomly collected from Omitara, Harnas and Otjinene. From each location, two healthy plants were collected annually and on the same site. Roots, stems, and leaves were separated and kept in different sterile plastic bags, totaling six samples each site. Plant samples were kept in a cooler box waiting to be transported to the laboratory where they were immediately processed. For the isolation and analysis of diazotroph, each plant was separately processed.

3.2 Endophytic bacteria isolations and growth conditions

Endophytic bacteria were isolated from roots, stems and leaves of *T. esculentum* plants from Omitara, Harnas and Otjinene. These samples were separately washed using running tap water to dislodge any soil or dust particles. It was followed by vigorous shaking with sterile water to remove any epiphytic microbes.

Samples were then immersed for 30 sec in 70% ethanol solution and afterwards immersed in 100 ml of 2% sodium hypochlorite containing 0.1% Tween 20. To remove the disinfectant, the roots were rinsed five times in two washes of RNase free sterile water. Roots, stems and leaves were dried on sterile paper towels. Negative controls were prepared. Water from the final rinse was plated onto Tryptic Soy Agar (TSA) to find out if the surface sterilization technique was successful. When no growth was observed, this confirmed that the surface-sterilization procedure was effective.

For growth of heterotrophic nitrogen-fixing bacteria, nitrogen-free Synthetic Malate (SM) modified medium was used for enrichment. At first, 1 g of root, stem or leaf from each sampling site was sliced with a sterile scalpel. To macerate samples, a sterile mortar and pestle to which a pinch of sterile sand were used. Macerated samples were incubated for a week at room temperature in nitrogen free SM medium (Reinhold, Hurek, Niemann, & Fendrik, 1986). The modified SM composition was per L: 1g of DL-malate, 1g of glucose, 1ml of ethanol, 20 mg of yeast extract and 1g of vitamins (Hurek, Reinhold-Hurek, Montagu, & Kellenberger, 1994). The one week incubation at room temperature was followed by two weeks incubation at 30°C. The tissue extracts from SM medium were serially diluted in phosphate buffer and plated in triplicate on VM-ethanol plates to recover any bacterial endophytes present in the plant tissue. Dilution series (up to 10^{-7}) were made and 0.1 ml aliquots were spread on VM-ethanol plates. All the plates were incubated at 30°C for 5–7 days. VM ethanol medium consisted of (per liter): K_2HPO_4 , 0.6 g; KH_2PO_4 , 0.4 g; NH_4Cl , 0.5 g; $MgSO_4 \cdot 7H_2O$, 0.2 g; $NaCl$, 1.1 g; $CaCl_2 \cdot 2H_2O$, 0.026 g; $MnSO_4 \cdot H_2O$, 0.01 g; $Na_2MoO_4 \cdot 2H_2O$, 0.002 g; Fe(III)-EDTA, 0.066 g; yeast extract, 1 g;

Tryptone, 3 g; pH 6±8; ethanol, 6 ml, sterilized by filtration (Reinhold-Hurek & Hurek, 2000). After selecting the most suitable plate (with colonies ranging between 30 and 300), the numbers of colony-forming units (CFU) were determined. Numbers of bacterial cells recovered were expressed as CFU g⁻¹ fresh tissue weight. Pure cultures were prepared by subculturing on VM-ethanol and incubated at 30°C. The bacterial colonies were initially screened and grouped by colony colour and morphological characteristics. The purified colonies were stored in a refrigerator at 4°C for further studies and for long-term storage at -20°C, the isolates were preserved in the glycerol stocks (20%).

3.3 DNA extraction and amplification of 16S ribosomal RNA gene

Genomic DNA was isolated from all the bacterial isolates and was used as template for PCR. Bacterial pure cultures in liquid Luria Bertani (LB) medium were grown on a shaker at room temperature (25°C) and 180 rpm/min for 32 hours. The cells were then concentrated by centrifugation at 4500 rpm for 2.5 minutes. The supernatant was discarded and the cells were washed three times with sterile double distilled water. Genomic DNA was extracted using the DNeasy Plant Mini Kit (QIAGEN, GmbH, Hilden, Germany) following the manufacturer's instructions.

Genomic DNA was prepared and was used as a template. 1.5 kb 16S rRNA genes were amplified using bacterial universal primers (Inqaba, Biotech, South Africa) 27F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-GGTTACCTTGTTACGACTT-3') (Weisburg, Barns, Pelletier, & Lane, 1991). The PCR mixture included 0.5 µM of each

primer, 23.5 µl of nuclease free water, 25 µl master mix and 1 µl of DNA template, making up a total of 50 µl. Similarly, the positive control contained all above and DNA template from *Pseudomonas putida* identified using 16S rDNA based analysis from previous studies. In the negative control, DNA template was substituted with nuclease free water.

Using the thermo-cycler (Bio-Rad, Hercules, CA), the PCR reaction was carried out as described: 1 cycle of pre-denaturation at 94°C for 4 minutes, 30 cycles of denaturation at 94°C for 1 minute, annealing at 50°C for 30seconds, Extension at 72°C for 1 minute, and with a final extension at 72°C for 10 minutes. Amplified PCR products were visualized on 1% agarose gel with a 1kb ladder (Inqaba Biotechnology Industries, SA) and further sequencing of PCR products was carried out to identify the particular isolates.

3.4 Detection of *nifH* gene by PCR amplification

To test for the ability for the bacteria to fix nitrogen template DNA was used to amplify approximately 400 bp *nifH* genes. Nested PCR was used as described by (Zhan & Sun, 2011). In the first PCR, primers, FGPH19 (5'-TACGGCAARGGTGGNATHG-3') and PolR (5'-ATS GCC ATC ATY TCR CCG GA-3') (Simonet et al., 1991; Poly, Monrozier, & Bally, 2001) were used. The PCR mixture included 12.5 µl master mix, 11.5 µl nuclear free water, 0.25 µM of each primer, and 0.5 µl of DNA template. The positive control contained all above and DNA template from *Enterobacter cloacae* identified using *nifH* gene based analysis from Haiyambo, Reinhold-Hurek, & Chimwamurombe (2015) . In the negative control, DNA template was substituted with nuclease free water.

In the thermo-cycler, the reaction was run under these conditions: pre-denaturation at 94°C for 4 minutes, denaturation at 94°C for 1 minute, annealing at 55°C for 1 minute, elongation at 72°C for 2 minutes, final elongation at 72°C for 5 minutes, for 30 cycles. In the second PCR reaction step, primers PolF (5' TGC GAY CCS AAR GCB GAC TC 3') and AQER (5' GAC GAT GTA GAT YTC CTG 3') (Poly et al., 2001) were used. The PCR mixture had a composition as the first PCR step, but 2 µl of the first PCR products was used as a template in this second step. PCR conditions of the first step were used for this (second) reaction (Sato et al., 2009). For quantification, both PCR products were viewed on 1% agarose along the DNA ladder.

3.5 16S rRNA and *nifH* sequences analysis

PCR products were sent to Inqaba Biotech for sequencing. When sequences were obtained, on both strands, a consensus was generated and edited using the BIOEDIT sequence alignment editor (Hall, 1999). Putative bacterial taxonomic affiliation of each library were assigned based on the closest match to sequence using the Classifier program of the Ribosomal Database Project (RDP) release 11 with confidence level of 80%. Phylogenetic trees were constructed using partial rRNA gene sequences of bacterial isolates obtained in this study. The resulting sequences were matched against those available in the GenBank database using BlastN algorithm (Altschul et al., 1997). They were then aligned using the multiple sequence alignment program CLUSTAL X (Tompson, Gibson, Plewniak, Jeanmougin, & Higgins, 1997) with default parameters.

The evolutionary history was inferred using the Maximum Parsimony method.

The bootstrap consensus tree inferred from 500 replicates was taken to represent the evolutionary history of the taxa analyzed (Felsenstein, 1985). Branches corresponding to partitions reproduced in less than 50% bootstrap replicates were collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches (Felsenstein, 1985). The Maximum Parsimony tree was obtained using the Subtree-Pruning-Regrafting (SPR) algorithm (Nei & Kumar, 2000) with search level 1 in which the initial trees were obtained by the random addition of sequences (10 replicates). The analysis involved 143 nucleotide sequences. There were a total of 667 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura, Stecher, Peterson, Filipski, & Kumar, 2013).

3.6 Statistical analysis

Analysis of Variance (ANOVA) was performed using PAST (Hammer, Harper, & Ryan, 2001) and SPSS 22.0 Softwares to detect difference among means of CFU of the entire sample collected from different locations and different tissues. The statistical significance of the differences between bacterial densities in samples from the different sampling location over 4 years were tested by two-way ANOVA followed by Tukey's significant difference test (Tukey's HSD, $p < 0.005$). For ecological association, the Principal Component Analysis was performed using PAST software (Hammer, Harper, & Ryan, 2001) to determine the interrelationship and specificity of endophytic bacteria among

different genera. The rarefaction curves were also obtained with the PAST software package to determine the adequacy of sequence coverage for each sample.

There is a need to reduce the dimensionality of a large data set of variable into a more manageable, smaller set of axis during analysis of these microbial community samples. Afterwards, their major variation patterns can be visualised. To that end, the principal component analysis (PCA) was applied (Gulumbe, Dikko, & Bello, 2013). It allowed analyses and comparisons of endophytic bacterial communities from *T. esculentum* leaves, stems and tubers. Principal Component finds hypothetical variables (components) accounting for as much as possible of the variance in a given multivariate data (Gergen & Harmanescu, 2012) . These new variables are linear combinations of the original variables. In this study, the variation between the principal components and microbial community samples were computed using the PAST software (Hammer, Harper, & Ryan, 2001). This method should help us to identify groups of variables (endophytic microbes) based on their loadings and groups of samples based on their respective scores.

In PCA, the eigenvectors determine the direction of maximum variability while the eigenvalues specify the variance. According to Schmit & Lodge (2005) the first scores axis explained the maximum amount of the variability in data set that can be explained by a single variable. The score of the second score axis explained the maximum amount of variation that was not explained by the first axis and so on for additional axes. PCA analysis provides loadings that show the contribution of the variable in the original data to the axis scores created by the analysis. This would allow finding out which species or

environmental factors in the original data set are the most important in distinguishing the sites or samples

The diversity and relationships among the bacterial endophyte populations obtained from different samples were evaluated using diversity indices. Calculating diversity indices needs similar sample sizes for all populations to be compared and an ideal index of diversity should consider both richness and evenness. Evenness is an estimate of the relative abundance of different genotypes making up the richness of a population. Simpson's index of diversity (S_i) and Shannon–Weaver index (H) are in general the measures of diversity that accounts for both richness and proportion of each species (Atlas & Bartha, 1998).

The most commonly used index of evenness is $E1$ that scales H by maximally expected number of genotypes (Grünwald, Goodwin, Milgroom, & Fry, 2003). We used the number of genera (S), Simpson's diversity index:

$$S_i = 1/D,$$

$$\text{Where } D = (\sum n(n-1) / N(N-1),$$

n is the total number of organisms of a particular species and N is the total number of organisms of all organisms of all species.

Shannon–Weaver index was calculated using the following equation:

$$H_s = -\sum (P_i \ln[P_i]),$$

and Evenness index ($E1 = H/[\ln S]$), where P_i is the number of a given genus divided by the total number of isolates observed.

The taxonomic affinity of endophytic bacteria occurring among the different parts of *T. esculentum* was described using the Jaccard's coefficient (JI). JI was used to measure the similarity between pairs of samples:

JI= $a/(a+b+c)$ where a is the number of species occurring in both samples, b represents the number of species restricted to sample 1, and c represent the number of species restricted to sample 2. JI ranges between 0 and 1. $JI = 0$ when there are no taxa shared and $JI=1$ when all taxa are shared.

3.7 DNA extraction and Pyrosequencing

Plant DNA was extracted using commercial Qiagen DNA Isolation Kit (QIAGEN, GmbH, Hilden, Germany) according to the manufacturer's instructions. The duplicate DNA extracts for each layer were pooled prior to downstream manipulation. The following universal 16S ribosomal RNA bacterial primer set 8F (5'GAGTTTGATCCTGGCTCAG 3') and 533R (5'TTACCGCGGCTGCTGGCAC3') covering V1–V3 regions of SSU were synthesized by Inqaba Biotechnology (SA) were used for the polymerase chain reaction (PCR).

Each 25 µl PCR reaction contained 12.5 µl master mix (Inqaba Biotechnology, South Africa), 11.5 µl nuclear free water (Inqaba Biotechnology, South Africa), 0.25 µl of

each primer, and 0.5 µl of DNA template. AC1000 thermal cycler (Bio-Rad, Hercules, CA) was used for the PCR according to the manufacturer's instructions.

To avoid any biased results or possible contamination, a negative control containing all PCR components in which the DNA template was omitted and substituted with RNase free water. After PCR amplification, the amplicons were visualized by gel electrophoresis (1% agarose) and revealed bands of the corresponding size. The PCR products were re-amplified by Inqaba Biotechnology (South Africa) for 454 pyrosequencing. Sequences analyses were performed as prescribed by the Ribosomal Database Project (RDP) (Cole et al., 2009).

3.8 Sequencing quality control and Bioinformatic analysis

Raw read data from the 454 pyrosequencing runs were processed through a Ribosomal Database Project (RDP) quality filter pipeline (<https://pyro.cme.msu.edu/init/form.spr>) (Cole et al., 2014). Through the initial process step, all reads that contained ambiguous nucleotides, contained a single nucleotide mismatch with the PCR primer or were of atypically length (<150bp or > 500bp) were removed. The tag and primers were trimmed off from the dataset using RDP initial process pipeline (Cole et al., 2014). Since the gene used is 16S, the orientation of sequences was checked. The presence of possible chimeric sequences was investigated using the CHIMERA_CHECK program of the RDP release 11.3. *T. esculentum* plastids and mitochondria 16S rRNA were identified using BLAST search similarity against the whole dataset. Those with 95% or more similarity to plant

sequences were discarded. The RDP classifier was used to assign bacterial taxonomy hierarchy. The remaining sequences were further aligned using the Complete Linkage Clustering. This tool serves to dereplicate sequences, calculate the distance between sequences and group sequences into clusters by the complete linkage clustering method (Cole, et al., 2014). The remaining sequences were subjected to complete linkage clustering using the pyrosequencing pipeline at RDP with a conservative 5% dissimilarity to define the OTUs.

CHAPTER 4: RESULTS

4.1 Isolation and enumeration of *T. esculentum* endophytic bacterial population

The bacterial endophytic population associated with leaves, stems and tuberous root of *T. esculentum* from Omitara, Harnas and Otjinene (Eastern Namibia) were successfully isolated. Population dynamics of the endophytic bacterial over 4 years are given in Table 1.

From 72 samples, 605 endophytic bacterial isolates were obtained. The number of colony-forming units per gram fresh weight (CFU) of culturable endophytic bacteria isolated from various plant organs of *T. esculentum* plants was determined (Table 1). The endophytic bacterial population ranged from a minimum of $(1.38 \pm 0.13)10^4$ CFU g⁻¹ fresh stem weight in 2013 in Harnas to a maximum of $(5.09 \pm 0.13)10^6$ CFU g⁻¹ fresh tuberous weight in Omitara in 2012 (Table 1).

In leaves, the maximum microbial density was recorded in Otjinene, $(3.73 \pm 0.24)10^5$ CFU g⁻¹ fresh leaf tissue weight in 2013. The minimum was $(6.3 \pm 0.97)10^3$ CFU g⁻¹ fresh leaf tissue weight in 2011 in Harnas (Table 1). In stem tissues, the microbial density ranged $(1.38 \pm 0.13)10^3$ – $(5.50 \pm 1.2)10^5$ g⁻¹ CFU fresh stem weight in Harnas in 2013 and 2014 respectively (Table 1).

In tuberous root tissue, our results showed that CFU were constantly higher than in stems and leaves. Both the highest root bacteria population density $(5.09 \pm 0.13)10^6$ g⁻¹

fresh root tissue weight and the lowest $(1.08 \pm 0.22)10^6$ g⁻¹ fresh root tissue weight originated from Omitara in 2012.

Colony forming units (CFU) were calculated from the viable culturable CFUs. CFUs were then converted to log CFU for easy processing and statistical analysis. This data was used to assess whether the bacterial density was significantly affected by the plant tissue (leaves, stems or tubers), the sampling location, and the sampling time or the interaction of these tree factors. After natural logarithm conversion, mean CFUs were normally distributed.

Table 1: Enumeration of putative endophytic diazotroph bacteria isolated from aseptically fresh *T. esculentum* tissue leaves (L), stems (S) and tuberous roots (T) (means of three replicates, $\pm$ standard deviation) from different locations (Omitara, Harnas and Otjinene) over a period of 4 years (2011-2014).

	2011	2012	2013	2014
	CFUg ⁻¹ fresh plant weight			
Omitara₁ L	(3.2 $\pm$ 0.7)10 ⁵	(2.9 $\pm$ 0.42)10 ⁵	(2.7 $\pm$ 0.22)10 ⁵	(2.40 $\pm$ 0.55)10 ⁵
S	(2.2 $\pm$ 0.6)10 ⁵	(1.6 $\pm$ 0.52)10 ⁵	(2.4 $\pm$ 0.54)10 ⁵	(1.04 $\pm$ 0.3)10 ⁵
T	(2.24 $\pm$ 0.32)10 ⁶	(1.08 $\pm$ 0.22)10 ⁶	(1.2 $\pm$ 0.45)10 ⁶	(1.5 $\pm$ 1.3)10 ⁶
Omitara₂ L	(2.4 $\pm$ 0.08)10 ⁵	(2.72 $\pm$ 0.26)10 ⁵	(1.66 $\pm$ 0.35)10 ⁵	(2.20 $\pm$ 0.25)10 ⁵
S	(1.51 $\pm$ 0.18)10 ⁵	(2.11 $\pm$ 0.15)10 ⁵	(2.26 $\pm$ 0.43)10 ⁵	(0.75 $\pm$ 0.3)10 ⁵
T	(3.57 $\pm$ 0.05)10 ⁶	(5.09 $\pm$ 0.13)10 ⁶	(1.87 $\pm$ 0.12)10 ⁶	(1.19 $\pm$ 2.7)10 ⁶
Harnas₁ L	(6.3 $\pm$ 0.97)10 ³	(2.54 $\pm$ 0.12)10 ⁴	(1.97 $\pm$ 0.92)10 ⁵	(1.65 $\pm$ 0.16)10 ⁵
S	(1.39 $\pm$ 0.11)10 ⁵	(2.28 $\pm$ 0.98)10 ⁵	(1.58 $\pm$ 0.17)10 ⁵	(5.50 $\pm$ 1.2)10 ⁵
T	(1.24 $\pm$ 0.07)10 ⁶	(1.73 $\pm$ 0.05)10 ⁶	(1.17 $\pm$ 0.11)10 ⁶	(1.88 $\pm$ 0.47)10 ⁶
Harnas₂ L	(2.04 $\pm$ 0.52)10 ⁵	(1.09 $\pm$ 0.13)10 ⁵	(2.49 $\pm$ 0.41)10 ⁵	(1.62 $\pm$ 0.35)10 ⁵
S	(2.51 $\pm$ 0.31)10 ⁵	(1.99 $\pm$ 0.11)10 ⁵	(1.38 $\pm$ 0.13)10 ⁴	(2.06 $\pm$ 0.72)10 ⁵
T	(2.02 $\pm$ 0.07)10 ⁶	(1.16 $\pm$ 0.04)10 ⁶	(3.02 $\pm$ 0.15)10 ⁶	(2.23 $\pm$ 0.42)10 ⁶
Otjinene₁ L	(1.29 $\pm$ 0.49)10 ⁵	(5.47 $\pm$ 0.14)10 ⁴	(3.73 $\pm$ 0.24)10 ⁵	(0.5 $\pm$ 0.11)10 ⁵
S	(1.22 $\pm$ 0.89)10 ⁵	(3.17 $\pm$ 0.18)10 ⁴	(2.34 $\pm$ 0.03)10 ⁴	3.2x10 ⁵
T	(1.26 $\pm$ 0.07)10 ⁶	(2.22 $\pm$ 0.14)10 ⁶	(1.97 $\pm$ 0.92)10 ⁶	(1.45 $\pm$ 0.42)10 ⁶
Otjinene₂ L	(2.37 $\pm$ 0.77)10 ³	(1.84 $\pm$ 0.12)10 ⁴	(1.57 $\pm$ 0.62)10 ⁵	(0.39 $\pm$ 0.03)x10 ⁵
S	(1.19 $\pm$ 0.16)10 ⁵	(2.78 $\pm$ 0.68)10 ⁵	(1.28 $\pm$ 0.37)10 ⁵	64x10 ⁴
T	(1.94 $\pm$ 0.27)10 ⁶	(1.43 $\pm$ 0.15)10 ⁶	(1.37 $\pm$ 0.21)10 ⁶	(1.85 $\pm$ 0.25)10 ⁶

Table 2: Two-way ANOVA showing that there is enough evidence to indicate that at least one of the three plant parts has a mean CFU different from other parts.

Dependent Variable: Mean CFU

Source	Type III Sum of Squares	df	Mean Square	F	Sign.
Location	83.807	2	41.903	1.086	0.348
Plant parts	4482.489	2	2241.245	58.108	0
Time	11.911	3	3.97	0.103	0.958
Location * Plant parts	110.923	4	27.731	0.719	0.585
Location * Time	240.985	6	40.164	1.041	0.415
Plant parts * Time	67.697	6	11.283	0.293	0.937
Location * Plant parts * Time	346.218	12	28.852	0.748	0.697
Error	1388.534	36	38.57		
Total	10679.676	72			

Table 3: Two-way ANOVA indicating that at 5% significance level at least one of the three plant parts has a mean CFU different from other parts (as p-value=0.000 is less than 0.05).

Tests of Between-Subjects Effects with Mean CFU were considered as independent

	Sum	of	Mean		
Plant parts	4482.489	2	2241.245	66.581	.000
Location	83.807	2	41.903	1.245	.295
Plant parts	4482.489	2	2241.245	66.581	.000
Time	11.911	3	3.970	.118	.949
Error	2154.357	64	33.662		
Total	10679.676	72			

Two-way ANOVA (Table 2) indicated that the time of sampling ($p > 0.05$), the location of plant samples ($p > 0.05$) and their respective interactions had no effect on the microbial density. However, it was indicated that at least one of the three plant parts had a mean CFU significantly different from others ($p\text{-value}=0.000 < 0.05$) (Table 2). Therefore, a pairwise comparison was required to identify which plant part mean CFU were responsible for rejecting the null hypothesis.

Table 4: Results of multiple comparisons performed on the total amount of bacterial cells to determine the plant part effect (1, 2, 3 represent respectively leaves, stem and tuberous root). Tukey test was used.

(I) Plant parts	(J)Plantparts	Mean Difference (I-J)	Std. Error	Sign.
1	2	-.352221	1.6748588	.976
	3	-16.911179*	1.6748588	.000
2	1	.352221	1.6748588	.976
	3	-16.558958*	1.6748588	.000
3	1	16.911179*	1.6748588	.000
	2	16.558958*	1.6748588	.000

It was revealed there is a significant difference microbial density between the above ground (stems and leaves) and belowground (tuberous roots) (Table 4).

4.2 Identification and Phylogenetic evaluation of cultured endophytic bacteria associated with *T. esculentum*.

Bacterial isolates were identified using their partial 16S rRNA gene sequences. They were compared to the Ribosomal Database Project (RDP) Classifier at 80% confidence threshold to establish the closest genus neighbours to respective type strains. It was indicated that all 605 sequences associated with *T.esculentum*'s leaves, stems and tubers were affiliated to three bacterial phyla: *Proteobacteria* (class alpha, beta and gamma), *Firmicutes* and *Actinobacteria*. Although the RDP classifier does not allow affiliating

bacterial isolate to a specific species, it had the advantage to assess the genus of most isolates. However, of the 605 isolates, 1 was an unclassified bacteria at the domain level (0.7%), while 4 *Bacillaceae* and 3 *Bacillale* were not classified. In the *Streptomycetacea* family 3 were unclassified. Four and 1 isolates from the *Xanthomonaceae* and *Pseudomonadaceae*, respectively were also unclassified. The most abundant isolates grouped as unclassified belong to the *Enterobacteriaceae* family where out of 291 isolates, 126 (43%) were unclassified.

The analysis of all sequences amplification products of the 16S rRNA genes led to the identification of 24 different bacterial genera namely: *Streptococcus*, *Lactococcus*, *Brevibacillus*, *Lysinibacillus*, *Paenibacillus*, *Bacillus*, *Arthrobacter*, *Curtobacterium*, *Streptomyces*, *Ochrobactrum*, *Rhizobium*, *Achromobacter*, *Burkholderia*, *Acinetobacter*, *Azomonas*, *Cosenzaea*, *Enterobacter*, *Escherichia/Shigella*, *Klebsiella*, *Kosakonia*, *Pantoea*, *Stenotrophomonas*, *Pseudomonas*, *Trabulsiella*, Unclassified *Enterobacteriaceae* Unclassified *Pseudomonadaceae*, Unclassified *Xanthomonadaceae*, Unclassified *Actinomycetales* and Unclassified *Bacilli*. The relative abundance of the different genus identified by RDP across the different samples is listed in Table 5. Most of isolates were related to common plant-associated, endophytic or soil bacteria. Major bacterial phyla distribution has been shown (Appendix 1).

Proteobacteria dominated the collection of sequences, comprising 67.4% of the total isolates with 0.7% of unclassified bacteria (Table 5). The sequences assigned to the phyla *Proteobacteria* included 3 classes *Alphaproteobacteria* (0.7%), *Betaproteobacteria*

(2%) and *Gammaproteobacteria* that was the most represented (97.3%). However, there were no isolates belonging to genera in the *Delta* or *Epsilon* class. In the *Gammaproteobacteria* sub-division the most abundant group of sequences was affiliated with the family of *Enterobacteriaceae* (73.4 %) followed by the *Moraxellaceae* (15.8%), *Xanthomonadaceae* (7.3%) and *Pseudomonaceae* (3.5%).

Table 5: Phylum and class distribution frequencies (%) of endophytic bacteria associated with *T. esculentum*'s tissues (leaves, stems and tuberous roots).

	Firmicutes	Actinobacteria	Proteobacteria		
			α -Proteobacteria	β -Proteobacteria	γ -Proteobacteria
Leaves	25(4.1%)	4(0.66%)	0	0	80(13.22%)
Stems	34(5.61%)	9(1.48%)	2(0.33%)	0	74(12.23%)
Tuberous roots	106(17.52%)	13(2.14%)	6(0.99%)	3(0.49%)	243(40.16%)

Firmicutes (23.7%) included the *Bacillaceae* family (83.9%), the *Planococcaceae* (11.2%) and one genera of the *Paenibacillaceae*. The genera *Bacillus* accounted for 97% of the *Bacillaceae*. *Actinobacteria* (4.3%), *Cyanobacteria* / Chloroplast (0.3%). All three phyla were found across the three sites and all samples. However, differences in class compositions were observed.

The highest number of sequences associated with leaves was *Gammaproteobacteria* (13.22%) and *Firmicutes* (4.1%). The *Firmicutes* were well represented in leaves, stems and tuberous root microbial communities, 4.1%, 5.61% and 17.5%, respectively. This was due to the fact the large number of the genus *Bacillus* and

Lysinibacillus are harbored in these tissues. A decrease in abundance was observed from tuberous root to the leaves. In contrast genera *Lactobacillus* and *Streptococcus* were only twice isolated in leaves. This high number was additionally found in stems with 12.33% and 5.61% for *Gammaproteobacteria* and *Firmicutes* respectively.

The *Proteobacteria* are prominent in tuberous roots. Their high number is mainly restricted to the class *Gammaproteobacteria* driven by the high abundance of the genus *Enterobacter* (13.74%) followed by the *Acinetobacter* (10.51%) of all bacteria occurring in the tuberous root. *Alpha* and *Beta proteobacteria* were poorly represented. These results demonstrated high bacteria diversity in leaves, stems and tuberous roots of *Tylosema esculentum*. The phylum *Actinobacteria* had the lowest diversity with only 3 genera represented namely *Arthrobacter*, *Curtobacterium*, *Streptomyces* and unclassified *Actinomycetales*. Half of them are found in the tuberous root (2.14% of the total microbial population recorded in this study). The second highest (1.48%) populated plant organ was the stem followed by the leaves (0.66%) (Table 5).

Table 6: Two -way ANOVA of the effect on the presence of bacterial phyla occurring in different *Tylosema esculentum*'s plant parts (leaf, stem and tuberous root).

ANOVA

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Plant Part	2755.167	2	1377.583	44.07287	2.1E-10	3.259446
Location	35851158	3	11950386	382327.4	4.27E-81	2.866266
Interaction Plant part*Location	925.1667	6	154.1944	4.933126	0.000896	2.363751
Within	1125.25	36	31.25694			
Total	35855963	47				

Each pair of test suggested that distinctly different tuber bacterial communities were present in the different locations and different part of the plants ($p < 0.001$) (Table 6).

The number of phyla in the three plant organs (leaves, stems and tuberous roots) though the same is characterized by a net dominance of the γ -*Proteobacteria* class in all plant parts. The four most frequent bacterial taxa (*Bacillus*, *Acinetobacter*, *Enterobacter* and the Unclassified *Enterobacteriaceae*) had a widespread presence in leaves, stems and tuberous root. Of the 24 genera isolated from *T. esculentum*, more than 50% of them did not occur in the leaves (*Streptococcus*, *Lactococcus*, *Brevibacillus*, *Paenibacillus*, *Arthrobacter*, *Curtobacterium*, *Ochrobactrum*, *Rhizobium*, *Achromobacter*, *Burkholderia*, *Azomonas*, *Cosenzaea*, *Escherichia/Shigella* and *Trabulsiella*). Similarly, the stems harboured fewer genera than the leaves. For example, isolates of *Brevibacillus*, *Arthrobacter*, *Rhizobium*, *Escherichia/Shigella* were not found in stems while, they were

present in leaves. The isolate of *Trabulsiella* was detected in stems, while it was absent in leaves. All recorded genera were found in the tuber with some, such as *Paenibacillus*, *Curtobacterium*, *Achromobacter*, *Azomonas*, *Cosenzaea* and *Trabulsiella* that were only recovered once in the tuberous root during the whole study period.

4.3 Community structures of endophytic bacteria associated with *T. esculentum*

Cluster analysis and diversity indices of Microbial communities

4.3.1 Cluster analyses

Based on similarities in bacteria genus assemblages, plant parts and the location, the dendrogram in Figure 3 has revealed two major clusters (Cluster I and II). However, plant parts from the same location did not always cluster together. In the Cluster I, all the tuberous root samples, year and origin distinctively clustered together. But, in Cluster II there were two distinct groups that shared approximately 57% similarity. In the same group, TOM2014 and TOT2014 clustered together and only share approximately 60% similarity. The remaining samples were similar and clustered together. The highest percentage similarity was 85% and was found between 2 tuber samples (TOM2014 and THARN 2013). The lowest similarity was 72% between (TOM 2013, THARN2013, TOTJ 2013) and (TOM2012, THARN2012, TOTJ2012, and TOTJ2011) (Figure3, Cluster I). In Cluster II, two distinct subgroups made of samples (TOM2013, THARN2013) were observed. The ANOVA (Table 6) conducted between the plant organ parts and their respective sampling locations supported this observation from the hierarchy dendrogram.

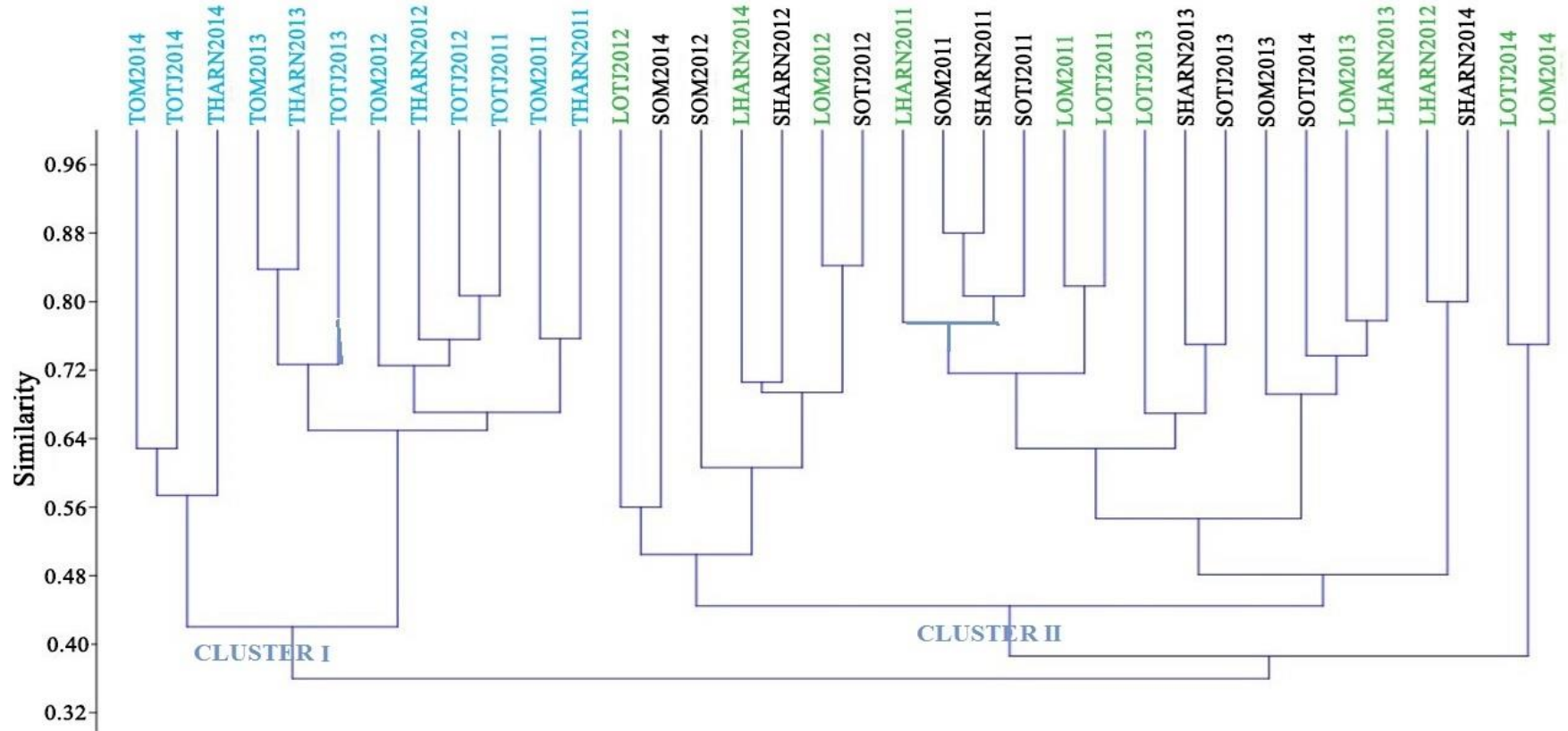


Figure 3: Hierarchical cluster dendrogram of bacterial communities' assemblages, plant parts and their respective locations. The sample designation shows plant parts by L, S, T as leaf, stem and tuberous root respectively; OM for Omitara, HARN for Harnas and OTJ for Otjinene and lastly the sampling year.

In Cluster II (Figure 3) of the Hierarchical cluster dendrogram of bacterial communities' assemblages, there was no clear cut clustering patterns about leaves and stems. Regardless of their origin, they were categorized in 9 different clusters. This clearly showed how, the above ground and below ground microbial communities might be different. In fact, based on the genus total occurrence analyses of variance test (Table 5), it was apparent that there was an interactive effect between microbial communities between plant parts and their location and even between their interactions ($p < 0.001$).

In Cluster II (Figure 3) of the dendrogram, some similarities were observed. Clusters such as (LOTJ2014, LOM2014) and (LOM2011, LOTJ2011) leave samples clustered as per their sampling date, while in clusters such as (LHARN2012, SOM2011, SHARN2011, SOTJ2011) and (SOM2013, SOTJ2013, LOM2013, LHARN2013, SHARN2013) despite being of different origin, they clustered according to year of sampling. This almost net separation of these two clusters can partly be attributed to occurrence of taxa observed in *T. esculentum*'s different plant organs (leaves, stems and tuberous roots).

4.3.2 Diversity indices

Estimates of diversity within (alpha) and between (beta) *T. esculentum* microbial communities were assessed for all 72 plant tissue (leaves, stems and tuberous root) samples in 2011-2014. These parameters varied depending on plant tissue and time. The

diversity indices of the microbial communities from different plant tissues of the host are listed in Appendix 3. The highest in both Shannon's diversity and richness indices of endophytic microbial communities were found mostly in the root tissues, while these two parameters of the endophytic bacterial communities were not constant across leaves and stems tissues (Appendix 3). Analysis based on the Shannon's index indicated that taxa diversity was low and did not vary among samples, sites and year of sampling (Table 6 and Figure 4). In this study, the Shannon's index ranged from the lowest 1.22 in leaves from Otjinene in 2013 (sample LOTJ2013) to 2.69 found in a tuber in Omitara in 2011 (TOM2011) (Appendix 3).

To determine if there is any significant difference between the three plant organs with regards to their Shannon's diversity index, evenness and similarities, normality of the data was tested using the Shapiro-Wilk test ($p= 0.9991, 0.7029, 0.2833$, respectively for the Shannon's index for leaves, stems and tuberous root).

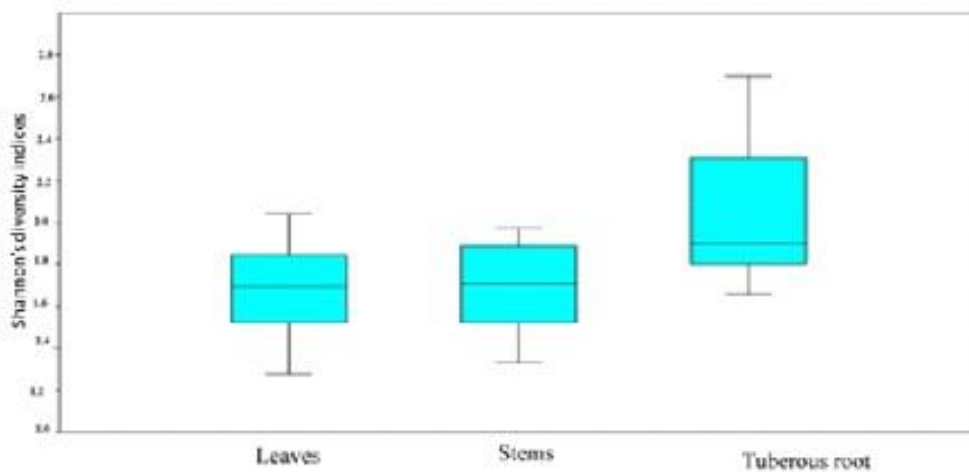


Figure 4: Boxplot that compared Shannon's diversity indices between the three different plants parts (leaves, stem and tuberous roots). Minimum and maximum values are indicated by the bars above boxes with the median values indicated by the inner blue box

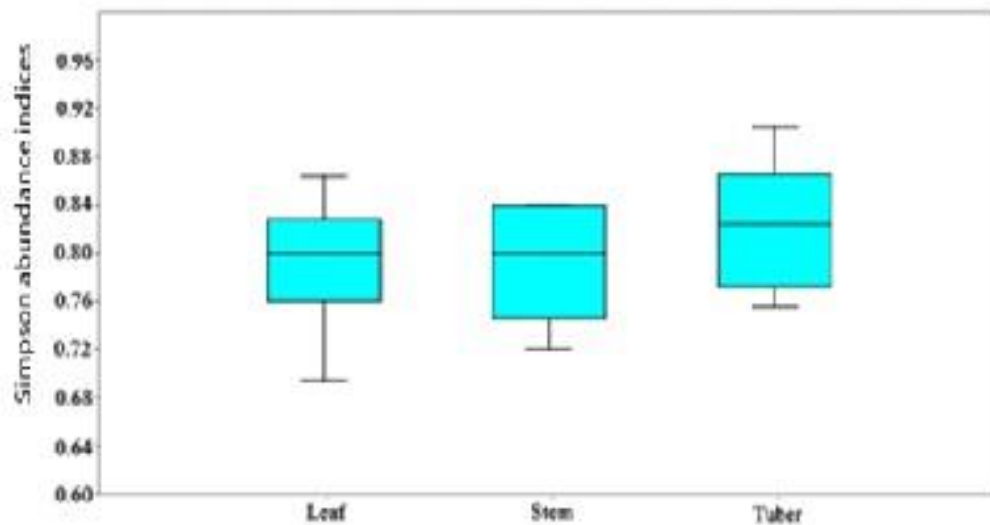


Figure 5: Boxplot that compared Simpson abundance values between the three different plant parts (leaves, stems and tuberous root). Minimum and maximum values are indicated by the bars above boxes with the median values indicated by the inner blue box.

Table 7: Two -way ANOVA of the effect on the presence of bacterial phyla occurring in different *Tylosema esculentum*'s plant parts (leaf, stem and tuberous **root**).

	Sum of sqrs	df	Mean square	F	p (same)
Between Shannon's diversity index:	0.881256	2	0.440628	7.508	0.002054
Within Shannon's diversity index:	1.93671	33	0.058688		
Total:	2.81796	35			

Tukey pairwise

	Shannon_HL	Shannon_HS	Shannon_HT
Shannon_HL		0.9947	0.005
Shannon_HS	0.1394		0.006
Shannon_HT	4.814	4.675	

In addition, based on Appendix 2 results, the Box plot of the Shannon's diversity index, and the one-way analysis of variance (ANOVA) and Tukey's *post hoc* (Table 7) tests, there is a significant difference between the tuberous root and leaves ($p = 0.005$) and stems' ($p = 0.006$) microbial communities (Table 7).

Simpson's diversity dominance index values were very low (Figure 5) and ranged between 0.72 (sample OM2012) and 1.0 in a tuberous root in Harnas 2012 (Appendix 3). It is higher in tuberous root compared to the leaves and stems. This would indicate that the microbial endophytic diversity is more in tubers than leaves and stems. When one dominant species coexists with a number of rare species, the condition is known as dominance.

Table 8: Test for equal means for dominance index

	Sum of squares	df	Mean square	F	Significance
Between Dominance indexes:	0.00697743	2	0.00348871	1.628	0.20
Within Dominance indexes:	0.0707138	33	0.00214284		
Total:	0.0776913	35			

The lowest Simpson dominance value was recorded in stems (0.72) in Omitara 2012,

while the highest Simpson dominance index (0.95) was in a tuberous root in Omitara in 2011 (Table 8).

4.4 Microbial community analysis of the three sites and plant parts using Principal Component Analysis (PCA).

To further investigate the effect of sampling location, plant part and sampling year on *T. esculentum*'s endophytic microbial community, a PCA was carried out.

Table 9: Principal component, Eigenvalues and percentage of explained variability generated after computing *T.esculentum* microbial samples data using PAST software (Hammer, Harper, & Ryan, 2001). Grey highlighted and bold are the extracted 4 PCs for this study with eigenvalues greater than 1.

PC	Eigenvalue	% variance	PC	Eigenvalue	% variance
1	11.8701	59.23	8	0.14202	0.71
2	3.01141	15.03	9	0.0551951	0.28
3	2.48184	12.40	10	0.0395548	0.20
4	1.31983	6.60	11	0.0272715	0.14
5	0.525212	2.62	12	0.0197499	0.10
6	0.335732	1.68	13	7.12E-33	3.55E-32
7	0.214163	1.07	14	2.85E-33	1.42E-32

Based on eigenvalues in Appendix 4 and on Ahrens (1982) and Gulumbe, Dikko and Bello (2013) assumption which stipulated that components with an eigenvalue higher than 1 should be retained, the original 29 PC were reduced to 4 main PCs (PC1, PC2, PC3 and PC4). Furthermore, these 4 PCs together have the highest PCA loadings that encompassed 93.2% of the total variance present in the original data (Table 9). The

decrease of 29 variables to 4 offers the advantage of making the reduced data manageable, while maintaining much of the variance in the data. The omitted components were of trivial importance to the total variance. They have eigenvalues of less than a unity (Appendix 4). According to Table 10, the first two principal components 1 and 2 accounted for 74.25% of variance and the remaining contributed to the residual approximately 25%.

According to Liu, Lin, & Kuo (2003), components loadings were classified as follows: the loading values greater than 0.75 signifies “strong”, the loading with absolute values between 0.75 and 0.50 indicate “moderate” while loading values between 0.50 and 0.30 designate “weak”. Using this classification and according to Figure 6 and Appendix 4, the positive loadings on the first component (PC1) (explaining approximately 59% of total variation) (Table 9) were moderate for *Bacillus* (0.74) and Unclassified *Enterobacteriaceae* (0.56) which reflects a moderate prevalence with the *T. esculentum* microbial community and weak for *Acinetobacter* (0.31). This indicates the moderate influence of these taxonomic groups on the observed variation, *Bacillus* in particular.

The second PC (PC2) (Figure 7) contributing to variation among the observed taxonomic groups represented approximately 15% of the variation (Table 9). Weak and negative loadings for this PC are represented by the taxa *Bacillus* (-0.33) and for *Acinetobacter* (-0.44) while the Unclassified *Enterobacteriaceae* and the genus *Enterobacter* have moderate loadings of 0.50 and 0.53 respectively (Appendix 4 and

Figure 11). This shows that there is a contrast between the occurrence of *Bacillus* and *Acinetobacter* and the *Enterobacteriaceae* family. There is a high *Enterobacteriaceae* and Unclassified *Enterobacteriaceae* and low *Bacillus* and *Acinetobacter* prevalence. The third component, PC3, explained 12.38% of the variance (Table 9) with loadings ranging from moderate (0.74) for *Enterobacter* to weak and negative (-0.46) for the unclassified *Enterobacteriaceae*. While the fourth component explained 6.5853% of the total variance (Table 9) and its loadings only for *Pantoea* (0.76) had a moderate one and others were weak and negative (*Lysinibacillus* (-0.39) and unclassified *Enterobacteriaceae* (-0.36) (Appendix 4)

4.4.1 Organ effect

In this Cartesian two dimensional spaces (Figure 8, A,B,C) of axes PC1 and PC2, in tubers, leaves and stems, four clear groups corresponding to the four years of sampling have been separated. This suggests, it may be possible to discriminate between microbial communities (Figure 8 C and D) that seem to overlap when analysed together (Figure 8 A). It can be seen that though clustering together the microbial community vary from year to year regardless of their geographical differences.

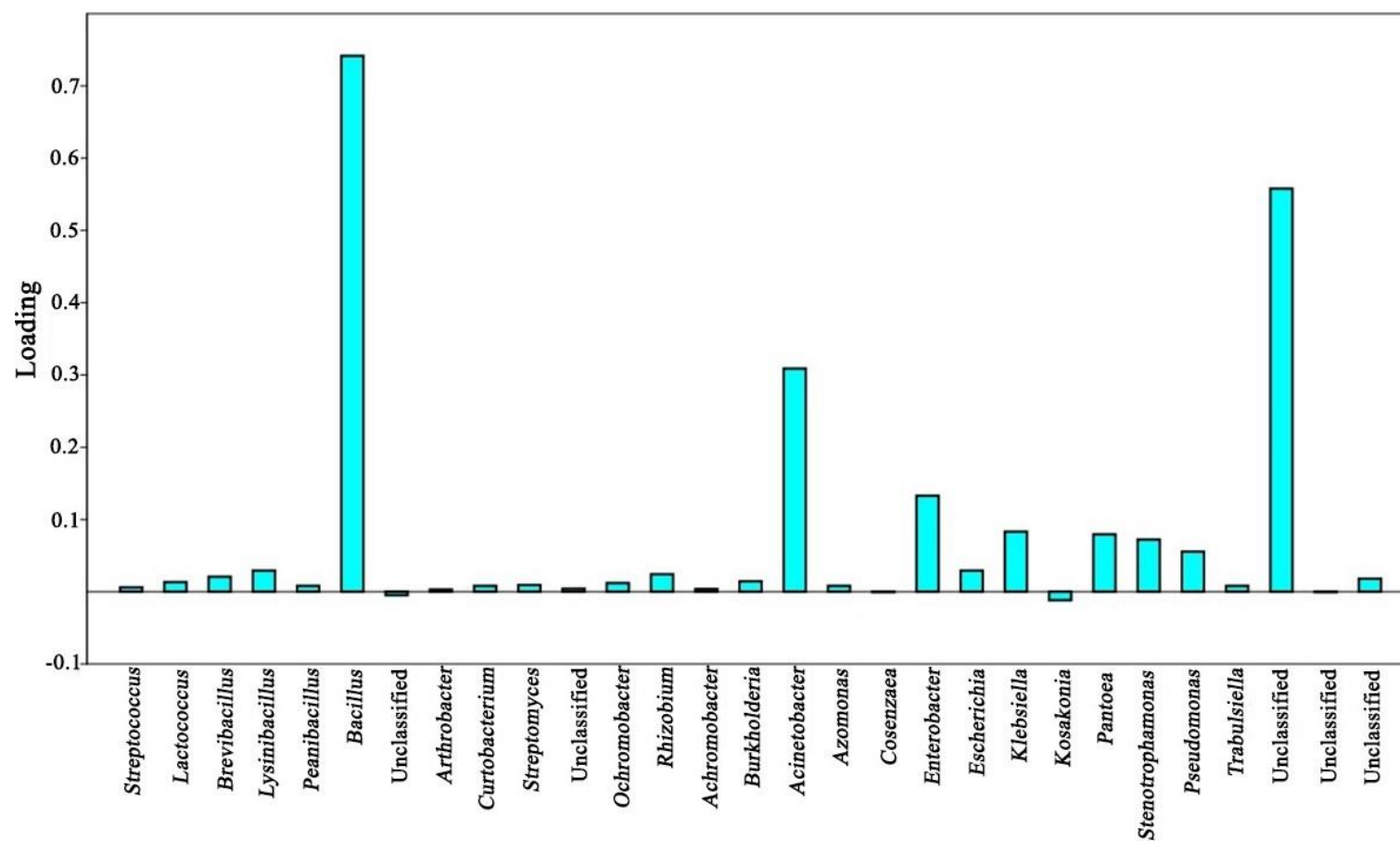


Figure 6: Loading of PC1 of community bacterial samples (microbial genus were used as variables in PCA)

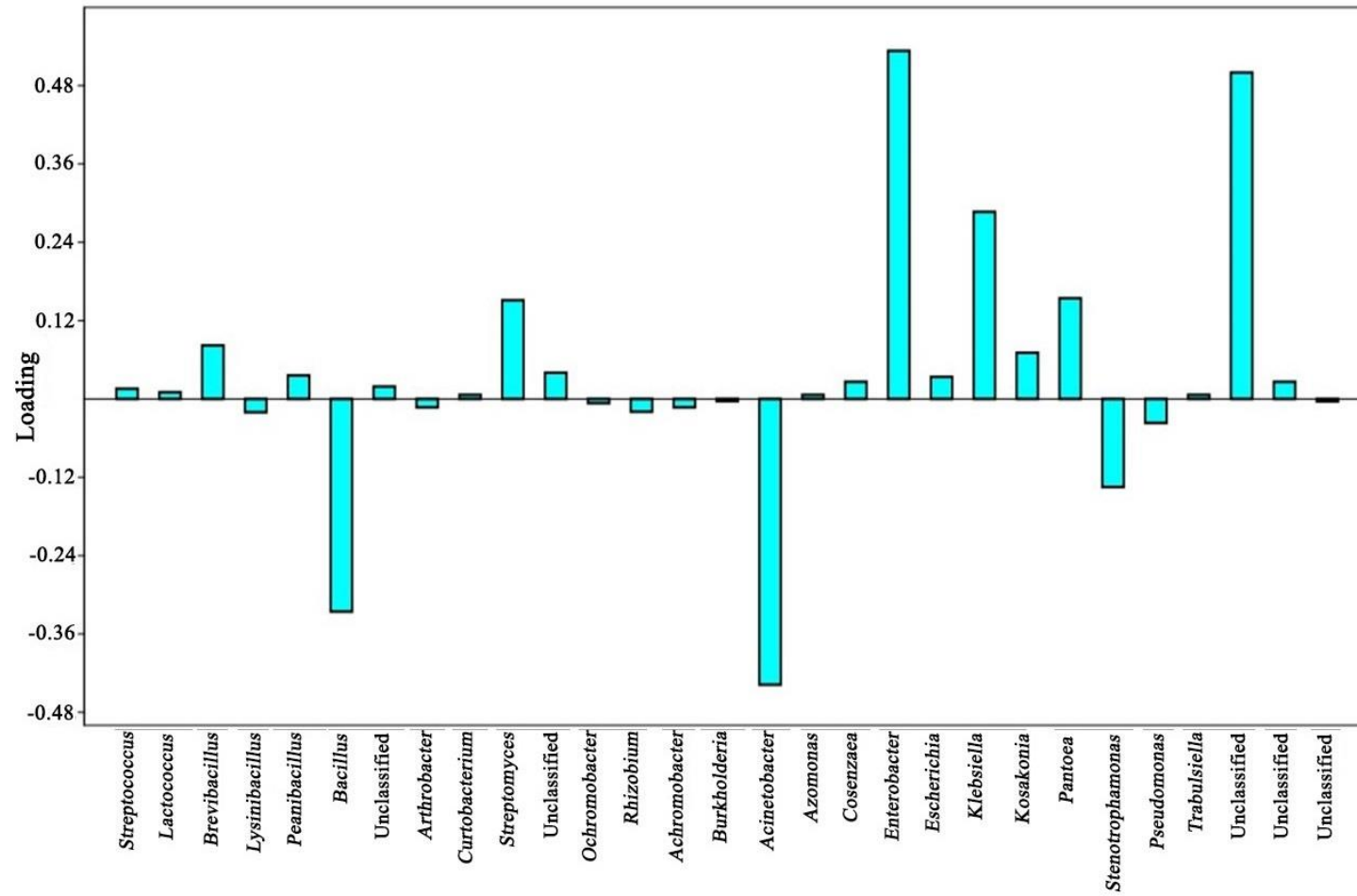
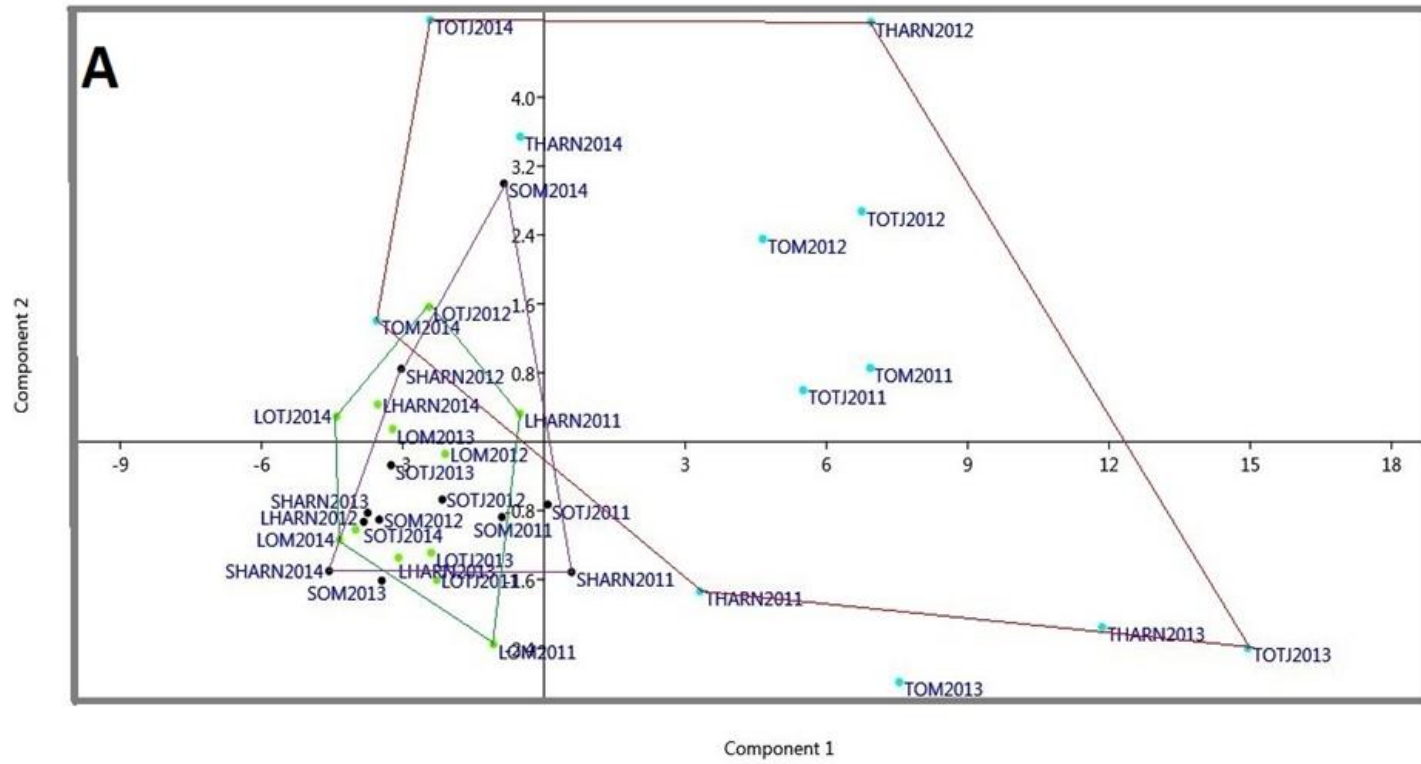
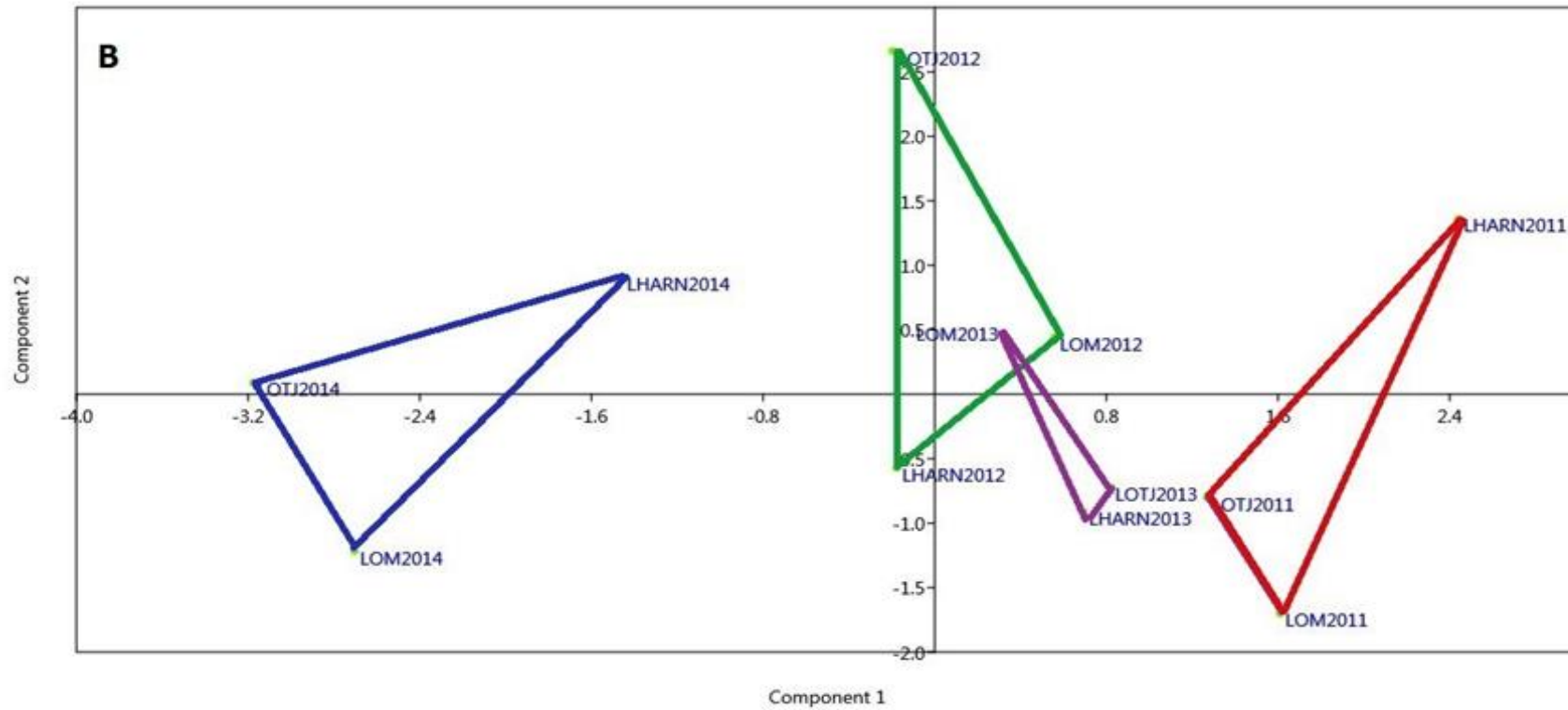
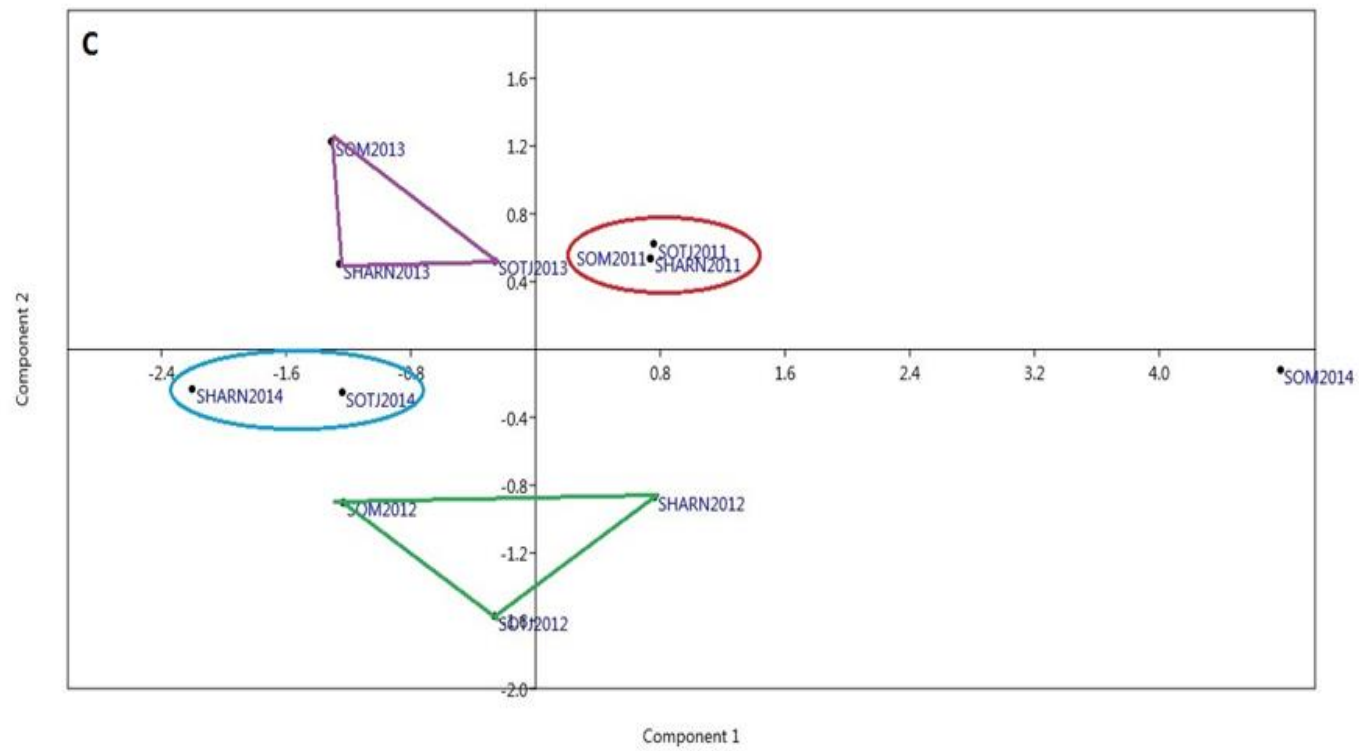


Figure 7: Loadings of PC2 of bacterial community bacterial samples (microbial genus were used as variables in PCA).







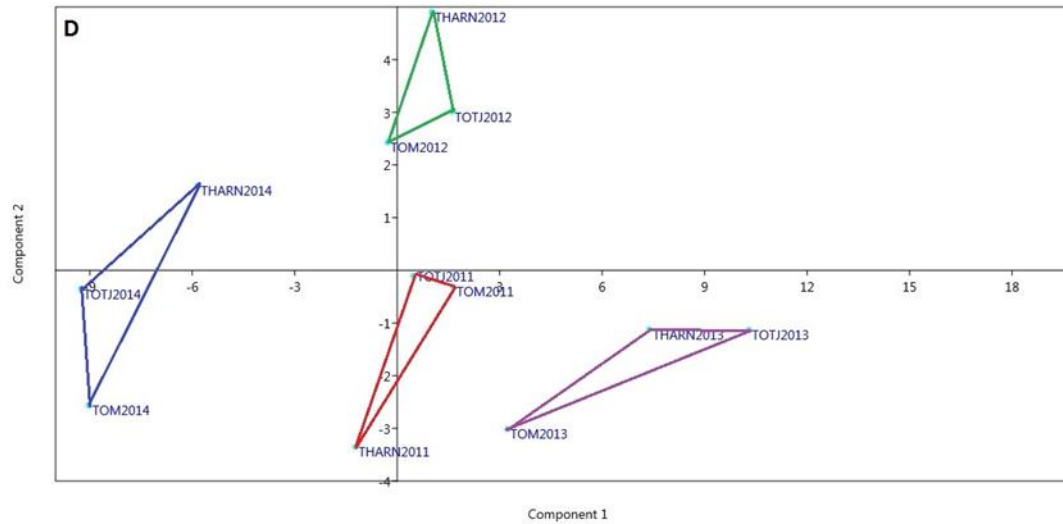


Figure 8: Clustering relationships between endophytic microbial communities based on principal component analysis (PCA). Here are shown figure (A): the total grouping of endophytic microbial communities in *T. esculentum*; (B), (C) and (D) separately clustered leaves, stems and tuberos root have shown a clear separation.

4.5 Phylogenetic analysis

4.5.1 Phylogenetic analysis of 16S-rRNA sequences

In order to investigate the genetic relationship between endophytic bacteria isolates associated with *T. esculentum*, their 16S rRNA sequences were analyzed with maximum parsimony to predict their diversity and assign them to molecular taxonomic units. Figure 9 shows the genetic relatedness of isolates from *T. esculentum*. Six clades were identified with a very strong bootstrap support of 99%, 99%, 91%, 88% and 89% respectively. These clades determined a high diversity and identity of endophytic bacteria that have been isolated from *T. esculentum*. The last clade was unresolved for it did not show any phylogenetic signal for 16S rRNA. This maximum parsimony tree that contained all sequences showed that endophytic bacteria associated with *T. esculentum* contained diverse bacteria.

In clade one, there were six classes with bootstraps ranging from moderate 55% and 65% to strong (80% to 99%) (Figure 9). Classes with moderate bootstrap values showed weak polytomic branching candidates that translate in a weak phylogenetic signal for the 16S rRNA gene used. More data would be needed to resolve their relationship and eventually their identities. Clade two though strongly supported with a high bootstrap value (99%) did not present deep branching, revealing the necessity of more data to resolve their identity and relationships (Figure 9).

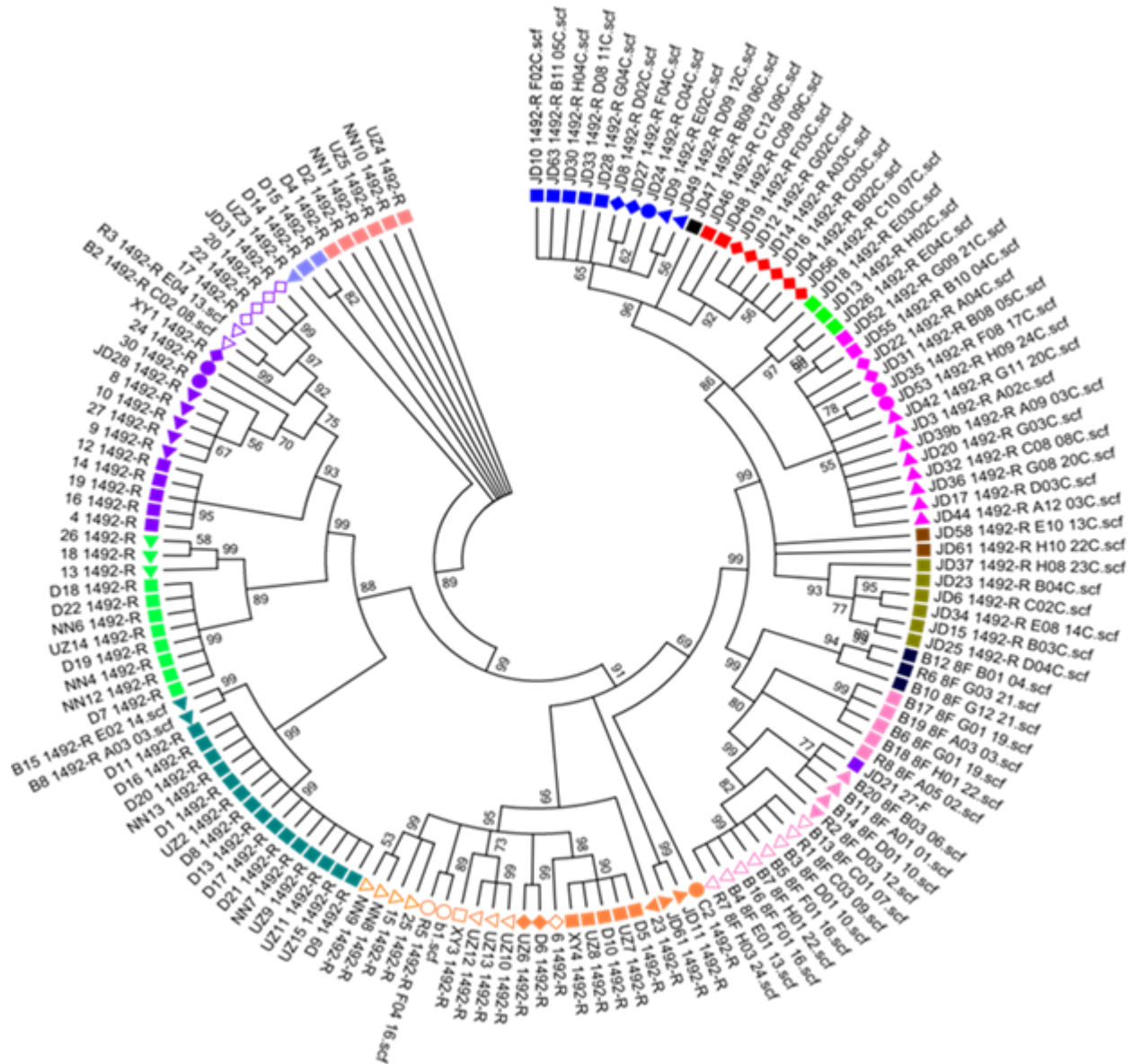


Figure 9: Strict consensus tree showing the phylogenetic relatedness of recovered endophytic bacteria associated with *T. esculentum* based on 16S rRNA sequences. The phylogenetic tree represents a maximum parsimony analysis of 143 taxa. Bootstrap values higher or equal to 50% (500 replicates) are shown at each branch.

Clade three consisted of an unresolved individual. Clade four had a strong bootstrap value (91%) with relatively deep branching attesting a more diverse clade.

Clade five was made of four classes of which only one presented a deep branching with a bootstrap value of 75%. The sixth clade was made of unresolved sequences (Figure 9). An additional phylogenetic analysis was performed (Figure 9) based on sequences associated with *T. esculentum* aligned along with their sister 16S rRNA reference strains with the highest similarity sequences obtained from the NCBI database. The results showed a tree of a similar typology as in Figure 10. With reference to RDP Database Classifier, and according to Stackebrandt and Ebers (2006) in Burbano, Gronmeyer, Hurek, & Reinhold-Hurek (2015), a great number of accounted taxa (49%) showed the identity similarity values below 98.7-99% (Table 10).

Clade one in Figure 10 was made of 9 distinctive classes. Clade one consisted of 51 sequences accounting for 36% of all endophytes isolated (Table 10). Blast matches for these sequences were close to members of the *Gammaproteobacteria* with 100% similarity index. A close examination of these results indicated that they are all members of the *Enterobacteriaceae* family. Of the 51 sequences, 33 sequences (JD10, JD63, JD28, JD33, JD30, JD8, JD27, JD24, JD9, JD49, JD47, JD14, JD56, JD16, JD4, JD12, JD46, JD48, JD19, JD21, B20, B14, B11, B2, R7, B5, B3, B4, B13, R1, B16, JD39B and B7) form the the majority and polytomic section of this clade. Only 30% of this clade's endophyte sequences had significant identities ($\geq 99\%$) to recorded species in the RDP database (Appendix 6). But the use of the 16S rRNA library could not clearly resolve the identity of these species.

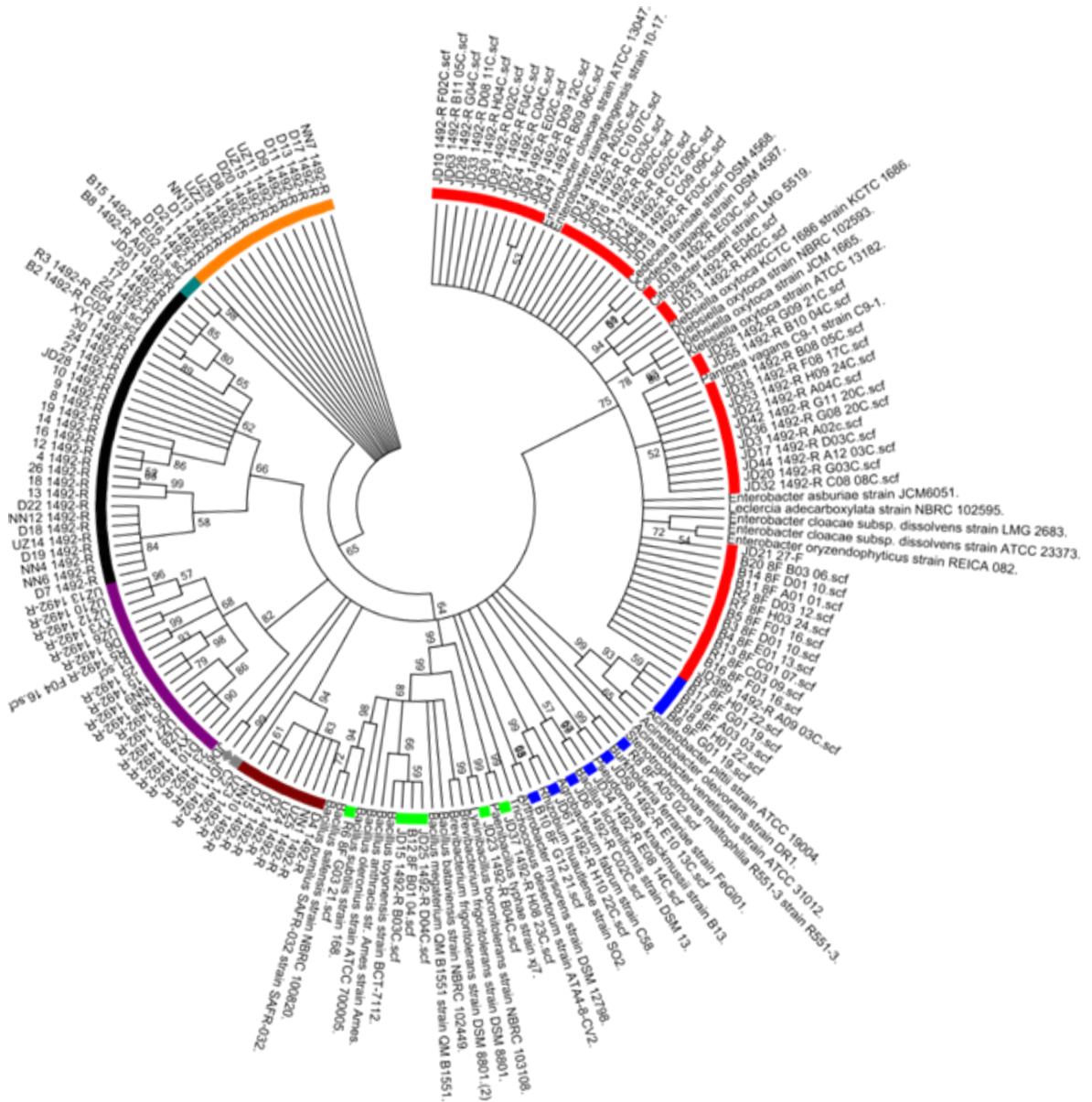


Figure 10: Strict consensus of the most parsimonious tree inferred using 16S rRNA sequences associated with *T. esculentum* based on 16S rRNA sequences. The phylogenetic tree represents a maximum parsimony analysis of 143 taxa with their reference strain. Bootstrap values higher than or equal to 50% (500 replicates) are shown at each branches.

Table 9: Comparison of culturable sequences associated with *T. esculentum*

Putative division	No. of total sequences (% representation)	% sequence similarity to its closest relative	No. of sequences that exhibit $\geq 99\%$ similarity to its closest relative
Clade I	51(36)	38-100	16 (31%)
Clade II	4 (3)	100	4 (100%)
Clade III	6 (4)	100	6 (100%)
Clade IV	6 (4)	67-100	4 (66%)
Clade V	9 (6)	47-100	8 (88%)
Clade VI	18 (13)	72-100	13 (72%)
Clade VII	29 (21)	39-100	4 (14%)
Clade VIII	2 (1)	100	2 (100%)
Clade IX	15 (11)	100	15 (100%)
Total	140		

This clade included the genera *Enterobacter* (percentage similarity 54-100%),

Pantoea (percentage similarity 87 to 100%), *Klebsiella* (67 – 80 % percentage similarity) and *Citrobacter* with the percentage similarity ranging between 42 and 62%. *Trabulsiella* and *Escherichia/Shigella* had both 38% percentage similarity to the closest strain from the RDP Database while *Raoultella*'s was a bit higher (48%) (Appendix 7). There was a weak phylogenetic support that would lead to a conclusion that 70% of these endophytes to a given genus. The remaining 30 % (JD18, JD26 and JD13) were nested with *Citrobacter* with a percentage similarity ranging from 87 to 100% and a bootstrap value of 94 and *Pantoea* nested with JD52, JD55, JD31, JD35, JD53, JD22, JD42, JD36, JD3, JD17, JD44, JD20 and JD32 with a moderate bootstrap value (52%) and 100% sequence similarity (Appendix 7).

Four sequences derived from the deeply branched clade two (B6, B17, B18, and B19) and accounting only for 3% (Table 9) of all sequences isolated from *T. esculentum* have been identified as *Acinetobacter* based on 16S rRNA sequence similarity (100%) of known taxa present in RDP database. A 99% bootstrap value for endophytes belonging to this clade was recorded. This attested that B6, B17, B18 and B19 all belong to the genus *Acinetobacter* (Appendix 6).

Clade three had five clusters (R8, JD5, JD34, JD61, and B10) (Figure 10) which accounted for 4% (Table 9). Their sequences showed 100% similarity with their best BLAST hits. This resulted in these clades being strongly supported with good bootstrap values ranging from 65 to 99% (Appendix 6). R8 (*Burkholderia*) and JD6 (*Bacillus*) sequence formed a distinct and moderately-supported clades with a bootstrap value of 65%

and 57% respectively. Their sequence similarity (100%) clades within *Betaproteobacteria* and *Firmicutes*. JD8, JD34, JD61, B10, JD37 and JD23 were all supported by a strong bootstrap value (99%) and a 100% sequence similarity clearly show they were respectively related to *Pseudomonas*, *Bacillus*, *Rhizobium*, *Paenibacillus* and *Lysinibacillus* (Appendix 6).

The taxonomic placement of endophytes belonging to Clade 4 showed that the similarity among the six sequences ranged between 67 and 100%. Sequences JD15, R6 and B12 were virtually identical displaying 100% sequence similarity. These 3 sequences belonged to 2 classes with JD14 and B12 sharing the upper class with a 99% bootstrap while R6 though sharing the same sequence similarity, showed a strong but lower bootstrap value (94%) suggesting all these sequences belong to the same genus *Bacillus*. JD25 sequence, however, though clustering with B12 and JD15 showed 67% sequence similarity (Appendix 6) with its best RDP database hit which resulted in phylogenetic placements with moderate bootstrap support. Accordingly, JD25 sequence is likely represented a new genus given its low sequence similarities (67%) with *Bacillus* strains. Finally, JD23 sequence was the only *Lysinibacillus* belonging to clade 4 with 100% sequence similarity and 99% bootstrap support and JD 37 was the only *Paenibacillus* of this clade. However, it had 91% similarity sequence with a strong bootstrap value suggesting the latter might represent a new genus. JD23 unequivocally represent the genus *Lysinibacillus*. This was backed by the high sequence similarity and strong bootstrap value (Figure 10).

In clade 5, samples sequences belonged to 3 classes (D4, NN1, UZ4, UZ5 D2; D14, D15 NN10 and UZ3). With the exception of UZ3 sequences which belongs to *Xylella* genus with 47% sequence similarity, other sequences (D4, NN1, UZ4, UZ5 D2; D14, D15 and NN10) had 100% sequence similarity to the RDP database and fell into *Xanthomonadaceae* family and *Stenotrophomonas* genus (Appendix 7). However, D4, NN1, UZ4, UZ5 and D2 sequences were supported with a strong bootstrap value (83%) while D14, D15 and NN10 showed a lower bootstrap value of 61%. With the exception of UZ3 despite its strong bootstrap value (94%) (Figure 9) which is most likely a new genus, other sequences probably belonged to the *Stenotrophomonas* genus.

The 18 sequences derived from clade 6 were divided into 6 classes that represented 18% of all selected sequences. Of all sequences of this clade 72% exhibited $\geq 99\%$ similarity to its closest relatives (Table 9). Despite their subdivision in different classes, all members of this clade belong to the *Bacillus* genus. They are all well supported with bootstrap values ranging from 57% for XY3 to 99% (D6 and UZ6) (Appendix 6). However, some sequences (XY4, b1, 25, UZ6) showed at least 87% sequence similarity with their best BLAST hit in RDP database which resulted in phylogenetic placements with good bootstrap support. They most likely represent a new genus given their low sequence similarities (87% at least and 98% at the most) with *Bacillaceae* type strains.

The 27 sequences that comprised the clade 7 were supported by a moderate bootstrap value (66%). They account for 21% of all sequences of which only 14% exhibit $\geq 99\%$ similarity to its closest relative. The current clade comprised 3 classes where the

third clade is made of 2 deep branching parts (4, 12, 16, 14, and 19) and (XY1, B2, R3, 22, 17, 20 JD31). These 2 classes are supported by a high to moderate bootstrap values, 86% and 65% respectively. All these sequences (4, 12, 16, 14, and 19) have sequence similarity below the accepted threshold. Like the previous group, these sequences (XY1, B2, R3, 22, 17, 20 JD31) did not cluster with any reference strain from the database. However, our phylogeny had established a close relationship of these species to the members of the genus *Enterobacter*, but due to their weak sequence similarity to the known database strains, their identity can not clearly be resolved to any particular *Enterobacteriaceae* (Appendix 6). Between the two deep branching classes, there was a more polytomic section of this class (Figure 10). The bootstrap value is moderate (65%) and their sequence similarity are far below the recommended threshold (Figure 9). As a result, there was not enough phylogenetic signals to conclude on their genus affiliation.

In phylogenetic studies, a 16S rRNA sequence identity between 98.5% and 99% is currently considered to be the absolute boundary for species restriction. Sequences (D7, NN6, NN4, D19, UZ14, D18, NN12, D22) and (13, 18, 26) constituted class one and two respectively. Despite their very close similarity in DNA sequence and phylogenetic affiliation to the *Enterobacter* and *Citrobacter* genus, there was no reference strain from the database that clustered with them. Taking all this into consideration, there is a weak support to confirm that these sequences might be belonging to these two genera of *Enterobacter* and *Citrobacter*.

Cluster 8 (dark green) (Figure 10) contained sequences B8 and B15 that clustered

with significant support. The bootstrap value was 98%, while the sequence similarity was 100%. These high values would indicate that these isolates belonged to the *Acinetobacter* genus (Appendix 7). However, these isolates never clustered with any reference strain from the database. In this case more data is needed to confirm the reliability of the isolates identification. Given the sequences obtained in this work were only mostly 700 – 800bp and were not perfect matches to the known type strains, it can be assumed that the bacterial isolates obtained in this study had unique strains Stackebrandt and Ebers, (2006) in Burbano, Gronmeyer, Hurek, & Reinhold-Hurek, (2015). There is a need to fully characterize these isolates in future.

4.5.2 Phylogenetic analysis of *nifH* sequences

The *nifH* encodes dinitrogenase reductase which is essential in nitrogen fixation in diazotroph organisms. It has been used as marker for the identification of nitrogen fixers (Rosch, Mergel, & Bothe, 2002). To confirm the potential for nitrogen fixation in endophyte housed in *T. esculentum* isolates, a nested PCR using degenerate primers as described by (Zhan & Sun, 2011) were used. This amplification yielded the expected 360-400 bp size product on agarose gel. Regardless of their origin, all isolated strains were positive for *nifH* gene amplification. This demonstrated not only the presence of endophytes but the putative nitrogen fixation ability of the microbial community associated with the non - nodulating *T. esculentum*. Thus, it has been proven incorrect that the presence of nodules is no longer a *sine qua non* condition for potential nitrogen fixation

in legumes (Reinhold-Hurek & Hurek, 2000).

The recovered *nifH* sequences ranged from 85 to 100% identity similarity to previously reported *nifH* protein sequences (Table 12 in Appendix 2). The phylogenetic positions of the representative *nifH* sequences are shown in Figures 11 and 12. At the first hit, it was established that 72% BLAST-N analyses of the 40 selected strains were closely related to *nifH* genes of uncultured bacteria, and the remaining 28% sequences were affiliated with the genera *Rhodovulum* (2 sequences), *Bacillus* (2 sequences), *Burkholderia* (2 sequences), *Azospirillum* (1 sequence), *Stenotrophomonas* (1 sequence), *Plectonema* (1 sequence) and one unidentified Bacterium (Table 12 in Appendix 2). Endophytic uncultured diazotrophs have been previously reported (Bürgmann, Meier, Bunge, Widmer, & Zeyer, 2005).

This suggested that they could be new unexplored species that could be of capital importance to the plant nitrogen input (Coelho et al., 2008). According to (Monteiro, Dutkiewicz, & Follows, 2011) since in the first hit, more than 50% of sequences matched most closely with uncultured organisms *nifH* genes, BLAST-N matching to a culturable strain was also considered (Table 12 in appendix 2).

In the second hit, the homology analysis of *nifH* sequences against NCBI databank sequence data showed that the majority of sequences were predominantly made of *Proteobacteria* (87.5 %). At the class level, most of *nifH* sequences showed highest similarities to sequences from *Alphaproteobacteria* (45 %), followed by *Betaproteobacteria* (40 %), *Gammaproteobacteria* (5%), *Bacillales* (7.5%) and *Cyanobacteria* (2.5 %). At the genera level (Figure 13), 14 genera were recorded.

In the second hit, the homology analysis of *nifH* sequences against NCBI databank sequence data showed that the majority of sequences were predominantly made of *Proteobacteria* (87.5 %). At the class level, most of *nifH* sequences showed highest similarities to sequences from *Alphaproteobacteria* (45 %), followed by *Betaproteobacteria* (40 %), *Gammaproteobacteria* (5%), *Bacillales* (7.5%) and *Cyanobacteria* (2.5 %). At the genera level (Figure 13), 14 genera were recorded.

Figure 12: Maximum parsimony tree of nifH gene sequence similarity, showing the position of nitrogen fixing isolates obtained from *T. esculentum* and their closest similar match from the NCBI database. Numbers, at the nodes show the level of bootstrap support based on 500 replicates, only values >50% were represented.

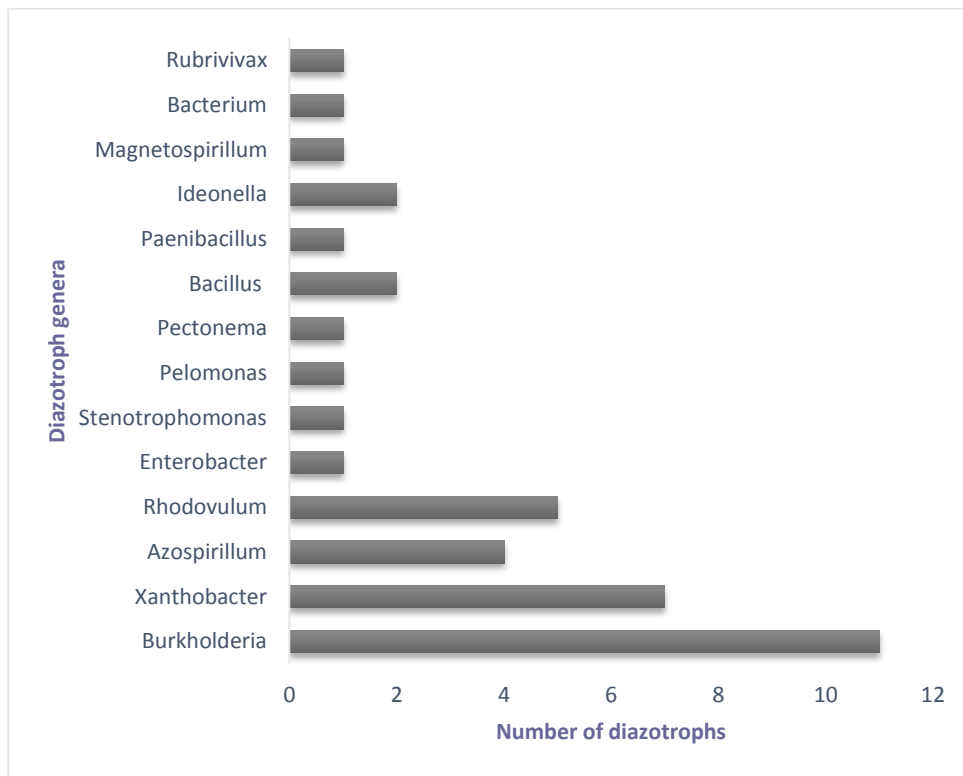


Figure 13: *NifH* distribution in 40 selected bacterial samples

The highest frequency of sequences was from *Burkholderia* (11sequences) followed by *Xanthobacter* (7sequences), *Rhodospirillum* (5sequences) and *Azospirillum* (4 sequences), *Bacillus* (2 sequences), *Ideonella* (2 sequences). *Enterobacter*, *Stenotrophomonas*, *Pelomonas*, *Pectonema*, *Paenibacillus*, *Magnetospirillum*, *Rubrivivax* and *Methylosinus* only one sequence of each was recorded (Figure 13, Table12 in Appendix 2). Phylogenetical trees (dendrograms) were constructed based upon the 16S rRNA gene sequences, all isolates were not closely related to reference strains. The reference strains shared a certain node but found far apart on different branches.

4.6 Comparative analysis of bacterial communities associated with different parts of *T. esculentum* as determined by Pyrosequencing.

4.6.1 Phylum, class-level analysis of bacterial communities in *T. esculentum* plant organs

In order to obtain the microbial community taxonomy hierarchy, all filtered sequences were subjected to the RDP pyrosequencing pipeline (Cole, et al., 2014). Default settings of the RDP pipeline were used, with a minimum length of 150 bp (Huse, Huber, Morrison, Sogin, & Welch, 2007). The naive Bayesian rRNA gene Classifier automatically estimates the classification reliability using bootstrapping. A subsample of 567 sequences which could not be assigned with bootstrap confidence was assigned as bacteria of uncertain affiliation.

From 10 plant samples, a total of 7620 raw reads with an average read length of 240 bp were processed following the Ribosomal Database Project (RDP) initial processing and the chimera removal (Cole, et al., 2014). The final number of reads passing quality filters was 3129 with an average read length of 294 bp. Of these, 567 (18%) sequences could not be classified at the phylum level according to the RDP Bayesian Sequence Classifier software and 358 (12%) were termed unclassified roots. 229 (7%) sequences were identified as *T. esculentum* chloroplast and discarded (Figure 14). The remaining 1970 (63%) sequence taxonomy was examined at the phylum level on the basis of the RDP Bayesian Classifier (Cole, et al., 2014) (Figure 14).

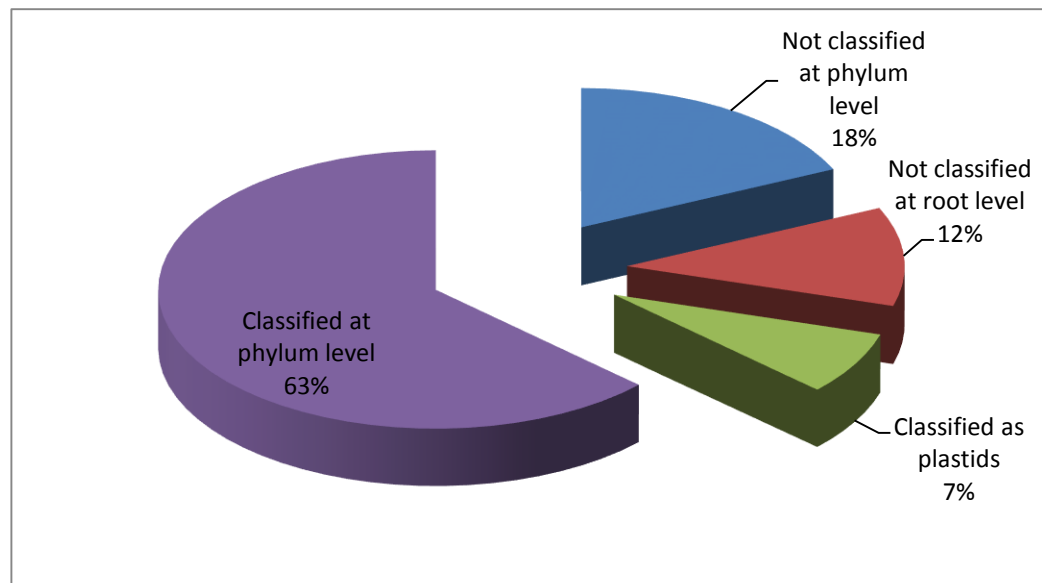


Figure 14: Bacteria phyla distribution recovered from *T. esculentum* using 454 sequencing methods.

All sequences could be classified into 5 phyla and 10 classes with the phylum *Firmicutes* (50.3%) affiliated sequences dominating followed by *Proteobacteria* (38.3%), *Fusobacteria* (5.7%), and *Actinobacteria* (4.5%) with *Bacteroidetes* representing only 1.4% (Figure 15, 16). Though present across all plant organs, *Firmicutes* were most dominant in tubers (samples 11T16S and 18T16S), while stems and leaves were shown to be highly dominated by the phylum *Proteobacteria* (Figure 15, 16, 17).

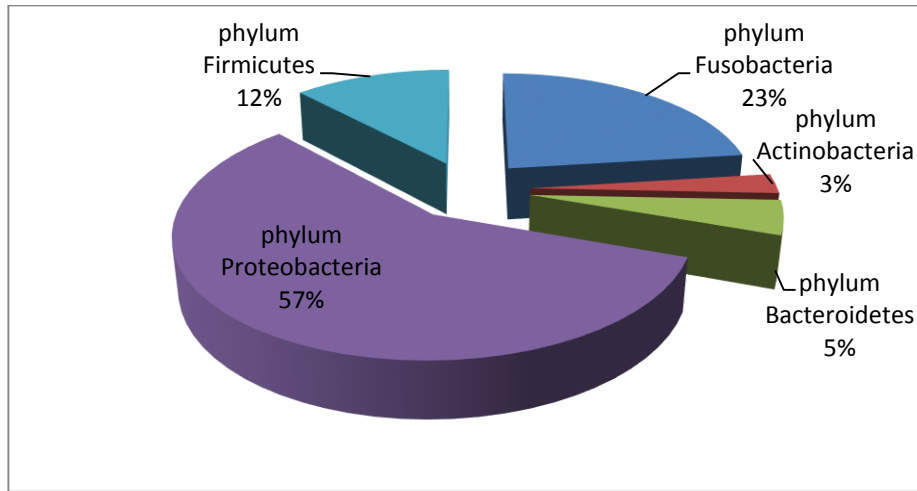


Figure 15: Above ground phyla distribution in *T. esculentum* as revealed by 454 sequencing approach.

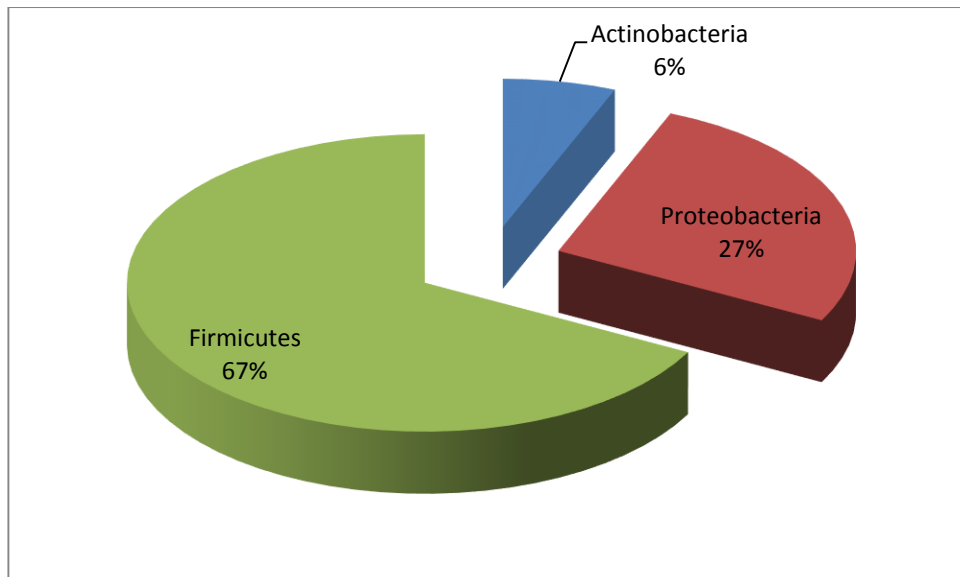


Figure 16: Below ground phyla distribution in *T. esculentum* revealed by 454 sequencing approach.

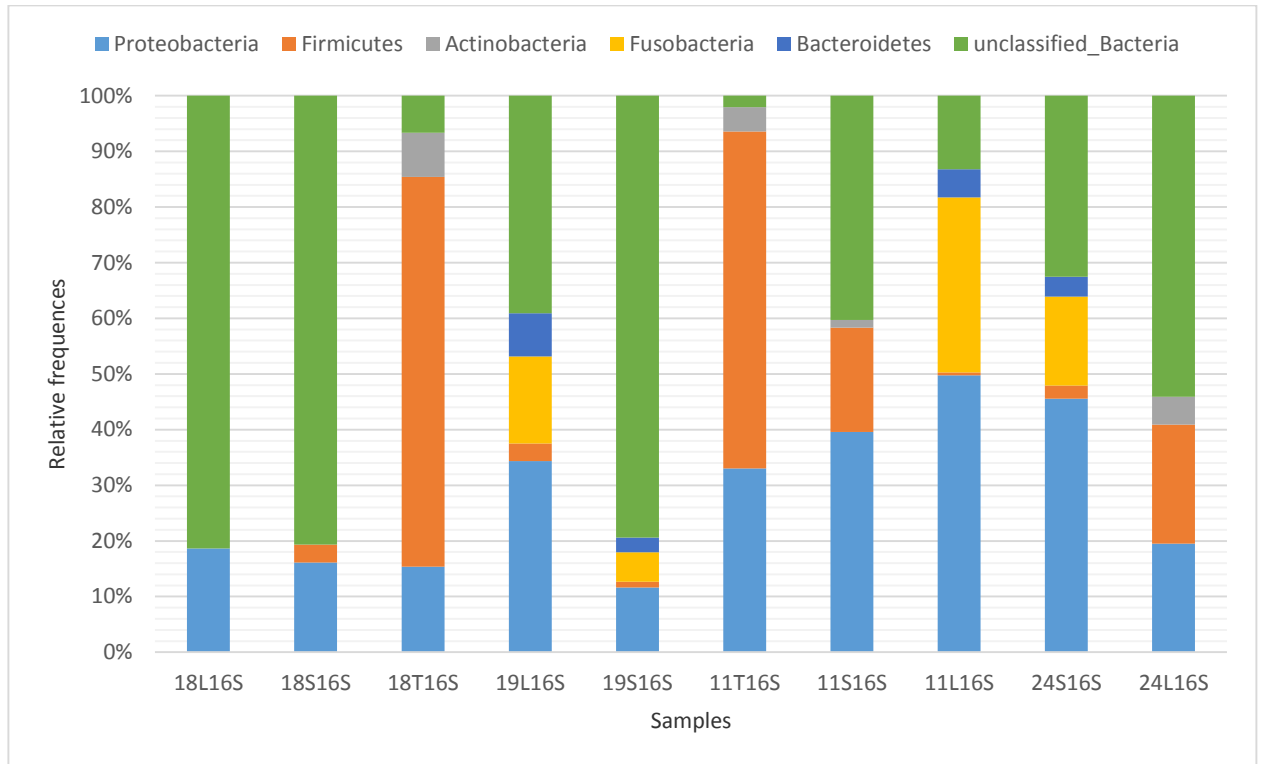


Figure 17: Phylum level abundance profiles using 16S rRNA sequence classification in *T. esculentum* bacteria community with respect to their origin (leaves, stems, and tubers). Columns represent the percentage of 16S rRNA sequences assigned to each phylum using the the RDP classifier (with a confidence threshold of 80%).

Firmicutes was the most abundant phylum in current study (Figure 17). It was highly represented in most of the samples with genera *Bacillus* affiliated sequences representing the bulk of samples accounting for the absolute majority ranging from 83.7% to 86% in tubers (Figure 17). The absolute majority of this class was mostly represented by 18T16S and 11T16S sample isolates (Figure 17).

The genus *Bacillus* represented one of the most diverse genera in the phylum *Firmicutes* (Garbeva, Van Veen, & Van Elsas, 2003) and was also a dominant

representative of *Firmicutes* in the current study. In leaf 24L16S and stem sample 11S16S, 56% and 59% of sequences were respectively assigned to the *Firmicutes* as well. Within this class there was a lower representation of the orders of *Lactobacillales*, *Clostridiales* with less than 10 sequences. Their respective genera were *Streptococcus*, *Parvimonas* and *Moryella*. *Parvimonas* that were represented by isolates occurring in leaves and stems of four samples (19L16S, 19L16S, 24L16S and 24S16S), while *Moryella* was only found in the sample 24S16S (Figure 17).

In this study, a large number (46.9%) of *Firmicutes* sequences were classified (Figure 17) to any existing family, order or genus. The biggest number (42.8%) of sequences to be classified to the order family in the *Bacilli* class was associated with tubers (18T16S and 11T16S) followed by the sample 24L16S with 24 sequences while the rest was represented by 6 and 1 sequences found in 11S16S and 18S16S respectively (Appendix 5). The second largest unclassified *Firmicutes* sequence group (3.78% sequences) belonged to the *Bacillaceae* family, whereas only one unclassified sequence belonged to the class *Bacilli* that occurred in the tuber (18T16S) (Figure 17).

Four (α , β , γ and ϵ) of the 5 classes of *Proteobacteria* were represented in this study. Of the 730 sequences affiliated to *Proteobacteria*, the largest fraction (82.5%) exhibited high similarity to *Gammaproteobacteria* and it was widely distributed in all samples (Figure 17, Appendix 5). The proportion of sequences that grouped with *Betaproteobacteria*, *Alphaproteobacteria* and *Epsilonproteobacteria* were 10%, 2% and 1.5% respectively (Appendix 5). The most predominant class, *Gammaproteobacteria*,

comprised of 5 genera namely *Pantoea*, *Leclercia*, *Enterobacter*, *Stenotrophomonas* and *Heamophilus* (Appendix 5). The latter was the most abundant. With the exception of the genus *Leclercia* whose sequence was once detected in the tuber, *Pantoea* and *Heamophilus* were found in leaves and stems. *Enterobacter* sequences were ubiquitous. More than 50% of *Gammaproteobacteria* sequences have been accounted as unclassified *Enterobacteriaceae*, unclassified *Xanthomonadaceae* and unclassified *Pseudomonadaceae* (Appendix 4).

These results suggested that the class *Gammaproteobacteria* might be a predominant inhabitant of *T.esclentum*. The second largest class was *Betaproteobacteria* with only 10% of all *Proteobacteria* sequences. They were related mostly to *Achromobacter* (48 sequences) and *Neisseria* (1 sequence) genus (Appendix 4 and Figure 17). Most *Alphaproteobacteria* belonged to two orders *Rhizobiales* and *Caulobacterales* represented by two genera *Brevundimonas* and *Ochrobactrum* found in tubers and leaves, but absent in stems (Appendix 4). Similarly, the class *Epsilonbacteria* and *Fusobacteria* were poorly represented and only found in tubers and leaves. This could be to due their low abundance or probably a low coverage. *Actinobacteria* were dominated by the genus *Streptomyces* with 48 sequences out of 85 of this phylum whereas 41 were not yet classified (Appendix 4).

Firmicutes showed generally an increase in relative abundance in the roots (Figure 17), especially the genus *Bacillus* in comparison to stems and leaves and the differences were substantially large (Appendix 5). Yet, they were consistently observed across

samples. However, the class *Clostridia* showed a different trend where it was exclusively found in leaves and stems.

The *Proteobacteria* phylum were the second largest and mostly dominated by the class *Gammaproteobacteria* (Figure 17). It was highly dominated by the *Enterobacteriaceae* and the unclassified *Enterobacteriaceae* (could only be identified at the family in the database). The *Pasteurellaceae* family did not shown major differences in relative abundance patterns. However, unlike the *Entrobacteriaceae* that was mostly found in tuberous root. The *Pasteurellaceae* family seemed to be leaf and stem dwellers. The class *Alpha* and *Betaproteobacteria* (*Achromobacter* and the unclassified *Alcalginaceae* family) were preferably found in tuberous roots, while the class *Epsilonproteobacteria* was found in tubers. *Actinobacteria* colonized all plant parts unlike *Fusobacteria* and *Bacteriodetes* that were found only on the above ground parts of the plant (Appendix 6).

4.7 Cluster analysis

Cluster analysis revealed that results between 20% and 30% coefficient threshold, the 10 samples were distinctively divided into two groups. The first group included 4 subgroups whose percentage similarities range between $\pm 38\%$ and 75%. Samples 19S, 16S and 19L16S showed 100% similarity which might originated from a labelling mistake from our sequencer provider company Inqaba.

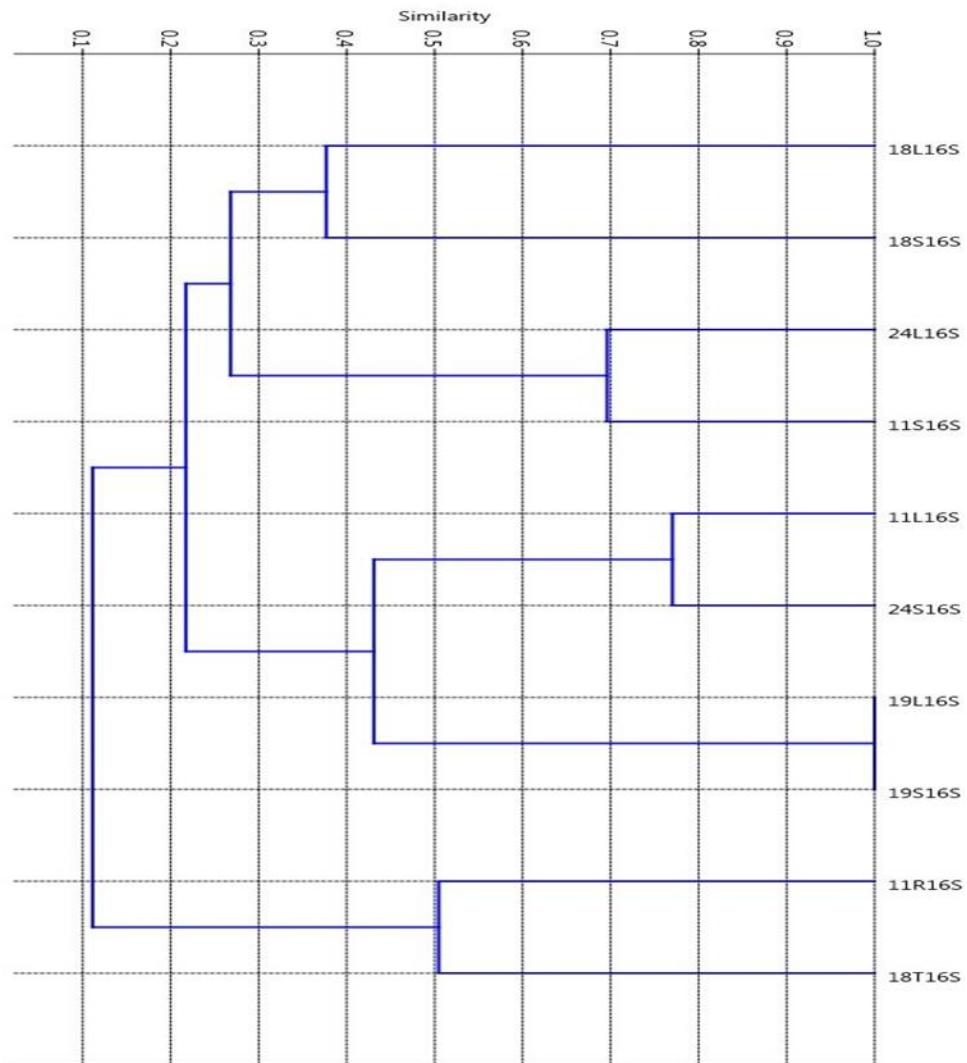


Figure 18: UPGMA dendrogram of 10 *T. esculentum* samples based on 16S rRNA gene sequences generated using Roche"454 pyrosequencing

The different samples clustered as follow: though samples 18L and 16L (leaf) and 18S and 16S (stem) were from the same plant and clustered together, their similarity index was however low (38%) (Figure 23). Other clusters are made of samples from different plants and parts such as 24L16S and 11S16S with the second highest similarity index (70%) and

11L16S and 24S16S that had the highest (75%) percentage similarity (Figure 23). Tuberos root samples clustered in a different group but they only share 50% similarity (Figure 18). This cluster analysis has shown that once again in *T. esculentum* there was a clear separation between the above ground plant parts (leaves and stems) and the below ground parts (tuberos roots).

4.8 Contributions of Bacterial Taxa to Principal Components

According to Figure 19 and the eigenvalues presented in the table 14, the PC1 accounts for 94%, whereas PC2 accounted only for 3.76% of the total variance. Hence, the dimensionality of the data set was reduced from 9 to 2. The PCA plot illustrated that there was a clear separation between the above and below ground parts of the plant *T. esculentum*.

PC1 (eigenvalue 223 343) accounted for 94% of the total variance and characterized endophytic bacteria of the phylum *Firmicutes* (Figure 20). In comparison to stems and leaves, this phylum showed a general increase in relative abundance in the roots especially the genus *Bacillus*. This explained the positive loading (0.42) in the first component (PC1) (explaining 94% of total variation) (Appendix 7) indicating the influence of this taxonomic group on the observed variations.

Above ground, the two communities of stem and leaf microbial communities were grouped together (Figure 19). According to PC1 and PC2 loading plots for leaves and stems only (Figures 22 and 23), microbes with major influences on the stems and leaves were termed “unclassified bacteria” and “unclassified root” (not yet identified at

root and bacteria level) with positive and negative loading coefficients (0.95 and 0.30) and (1.00) and 0.30) respectively. However and surprisingly, the unclassified bacteria and the unclassified root microbes had a great impact on the microbial communities in leaves and stems when dealt with separately.

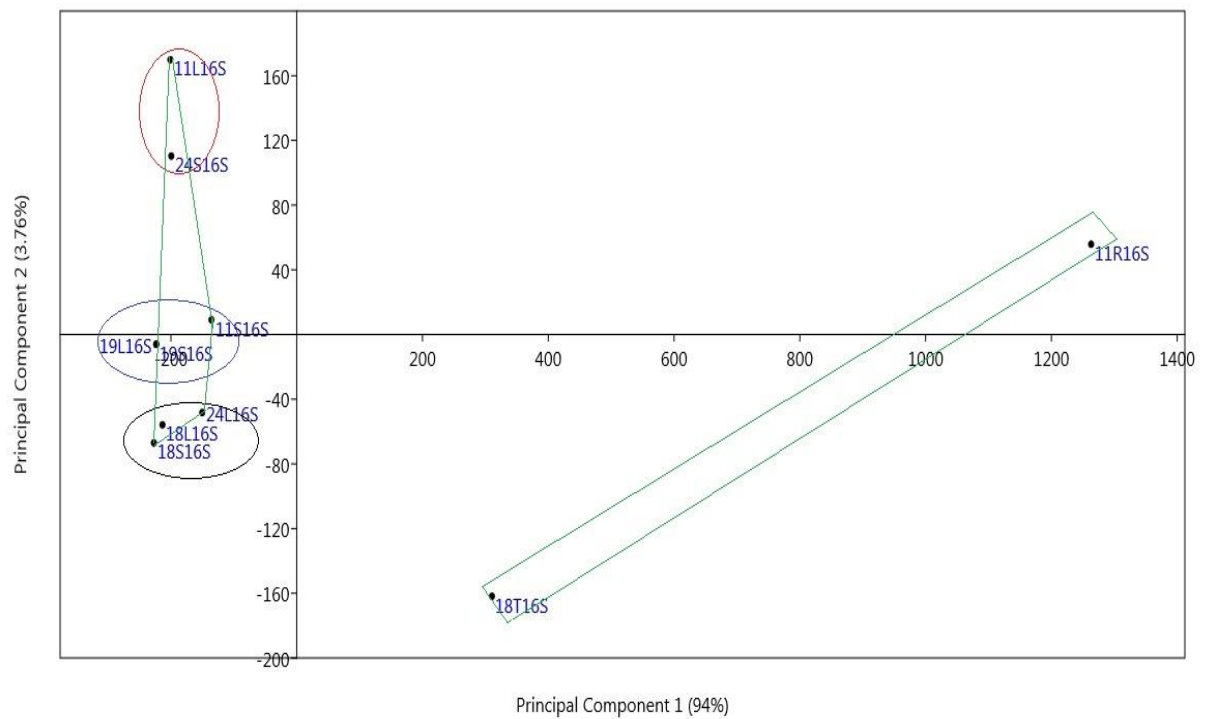


Figure 19: PCA ordination diagram illustrating the relationship among plots of PC1 versus PC2 scores of *T. esculentum* microbial communities distinguishing samples above (leaves and stems) and belowground (tuberous tuber).

PC2 (eigenvalue 8938.6) accounted for 3.80% of the total variance (Table 14). This PC 2 separated the class *Gammaproteobacteria* mainly consisting of the family *Enterobacteriaceae*

(Figure 17) whether identified or not. They were harbored in plant parts above ground as well as below ground with the highest diversity index below ground (Table 13). However, with the exception of the *Firmicutes* and most *Enterobacteriaceae* showed, in this study, to preferably occupy the tuber. It was not clear whether the *Firmicutes* and *Gammaproteobacteria* major influence and presence was due to ecological preferences of the endophytic bacteria or their host.

Comparing results from the Jaccard similarity matrix (Table 15), Figure 15, 16, 17 and 18 and the cluster analysis dendrogram and based on taxa's relative occurrence, there was a clear separation between *T. esculentum*'s above and below ground microbial communities. Moreover, the low Jaccard similarity (Table 15) between below ground and above ground indicated the low turnover between the tuber and the above ground microbial communities. It has just been shown that *Firmicutes* and *Proteobacteria* constituted groups of bacteria with major influence on *T. esculentum* microbial communities. Only the very low represented, *Clostridiales* and *Pasteurellaceae* had indicated to stay in leaves and stem only with no significant influence. An analysis of the particular above ground organ part to better comprehend the microbial assemblage revealed an astonishing result.

According to PC1 and PC2 loading plots for leaves and stems only (Figure 22 and 23), microbes with major influences on the stems and leaves were termed “unclassified bacteria” and “unclassified root”. This implies that these bacteria are not yet identified at root and bacteria level. This indicated that microbes occurring and having a great influence in the tuberous root were not closely related to ones found in the stems and leaves.

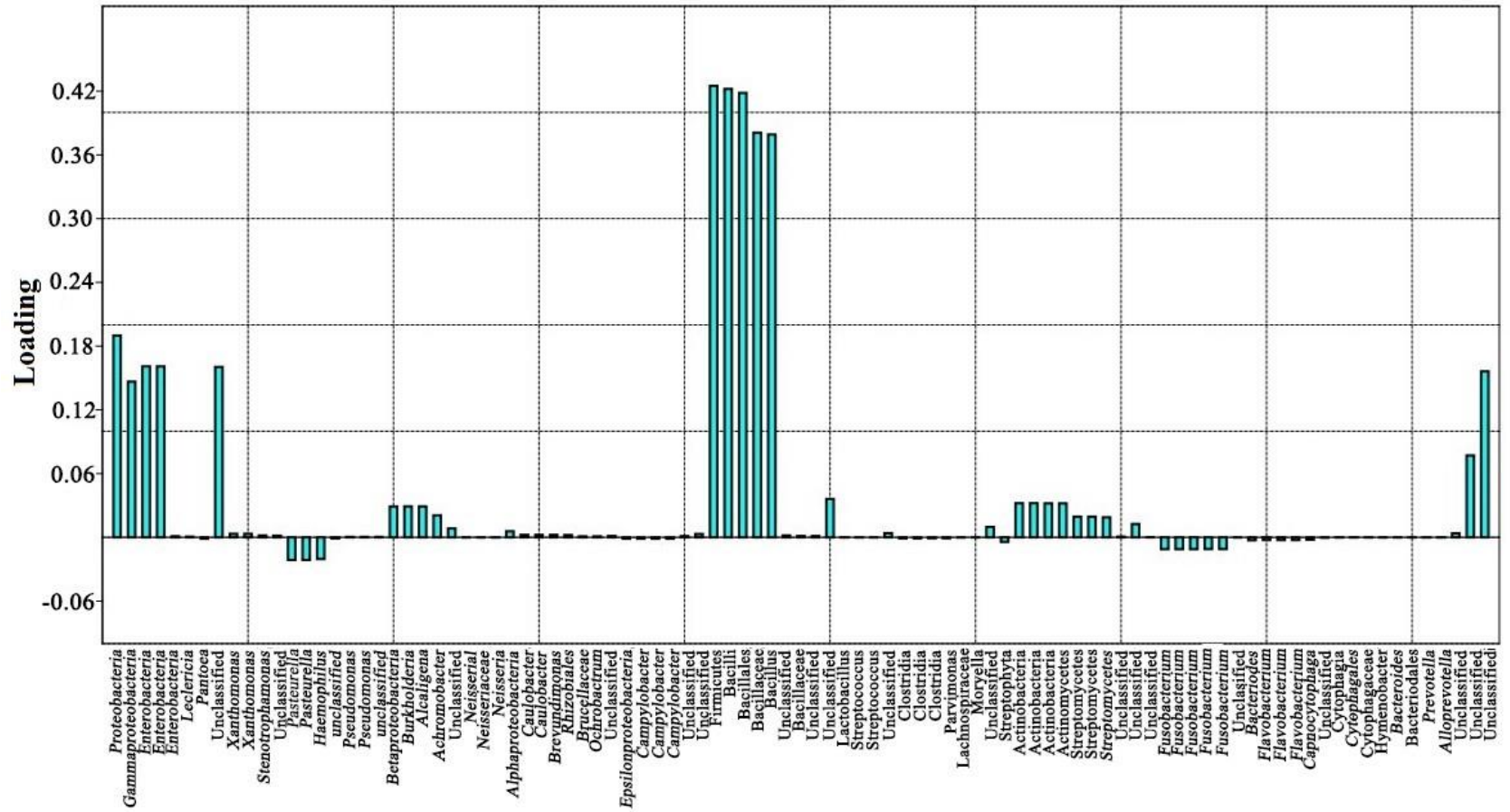


Figure 20: Loading plot for PC1 for total *T. esculentum* microbial community based on 16S rRNA gene sequences generated using Roche/454 pyrosequencing.

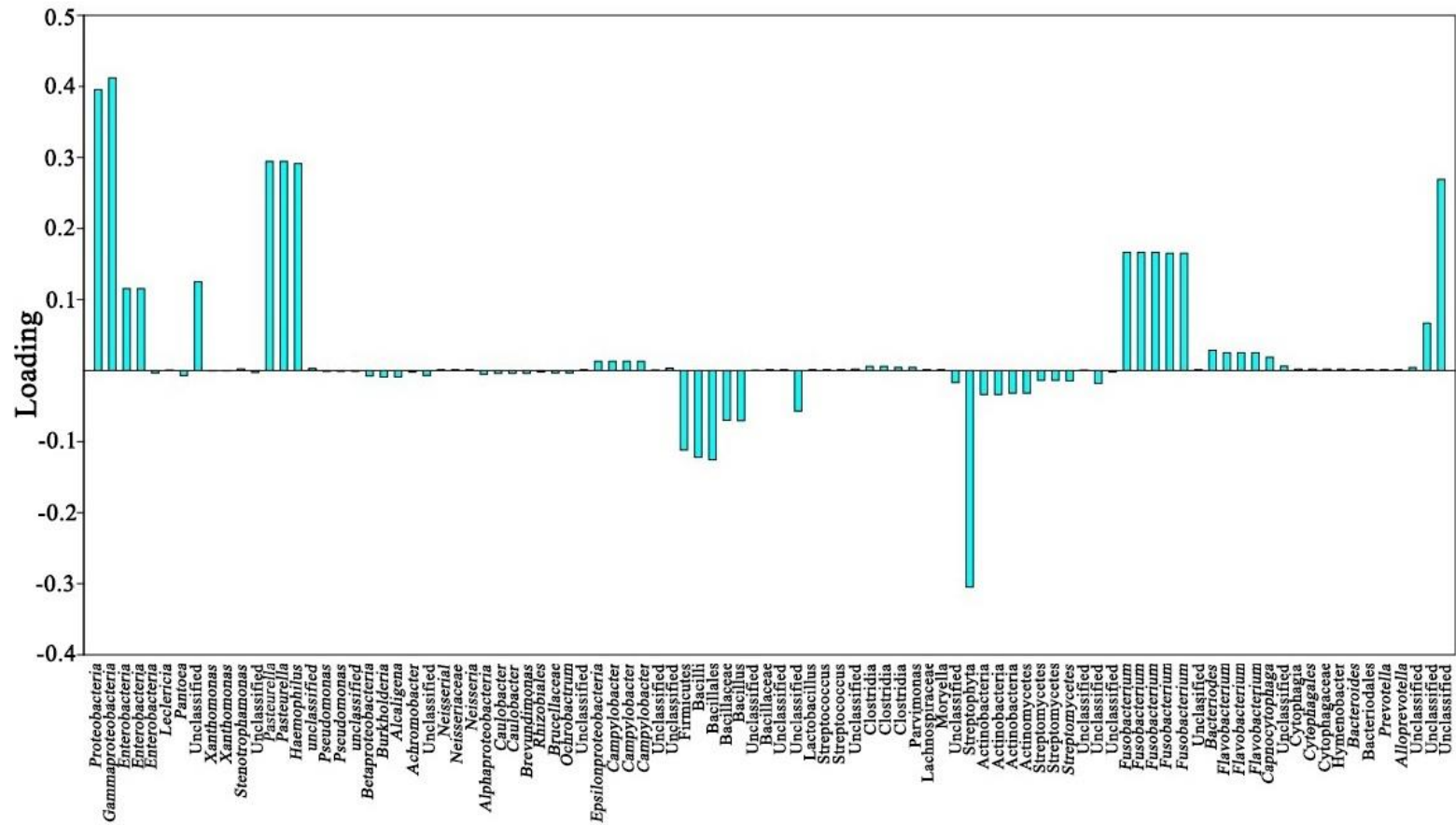


Figure 21: Loading plot for PC 2 for total *T. esculentum* microbial community based on 16S rRNA gene sequences generated using Roche/454 pyrosequencing.

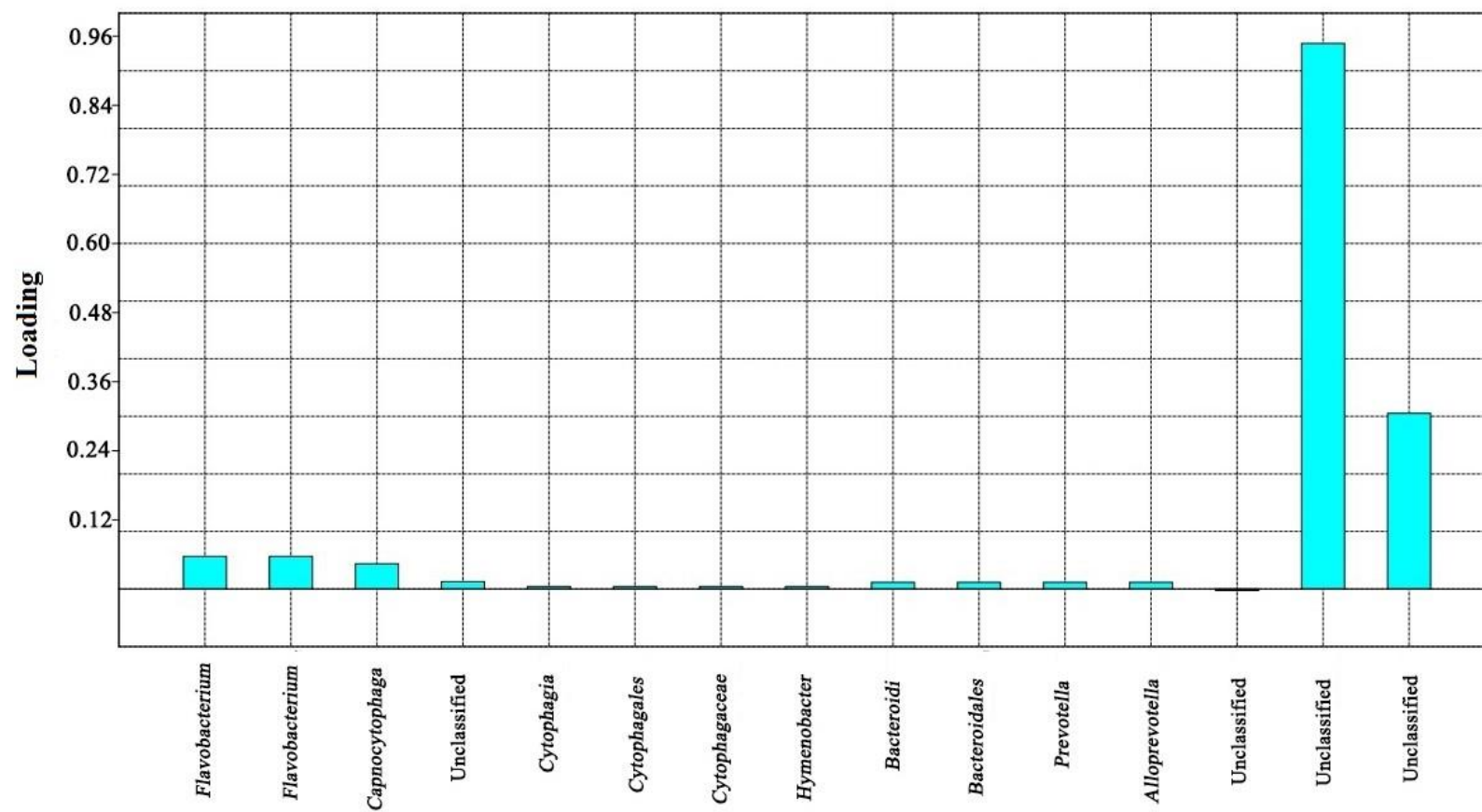


Figure 22: Loading plot for PC1 for leaves and stems *T. esculentum* microbial community.

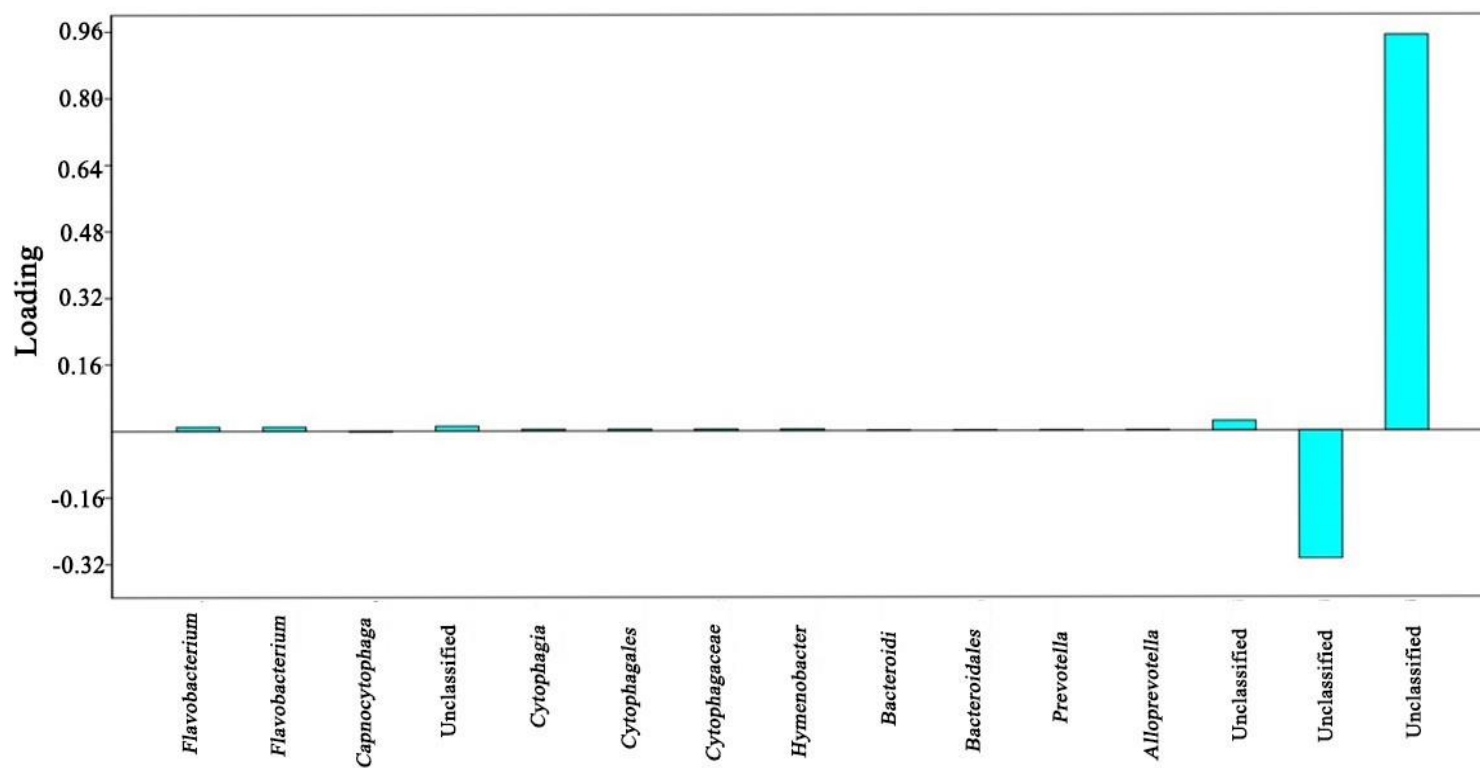


Figure 23: Loading plot for PC2 for leaves and stems *T. esculentum* microbial community.

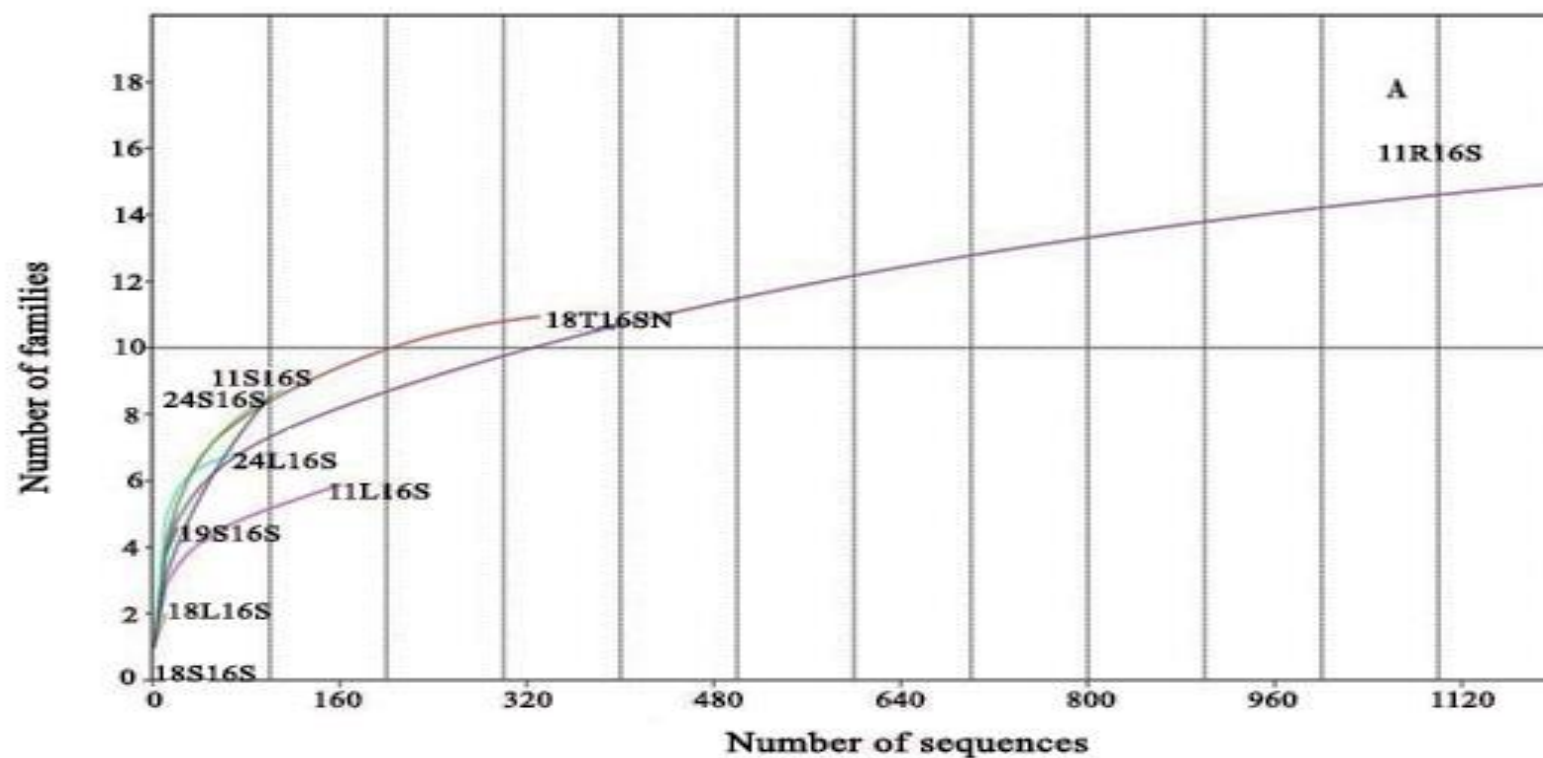


Figure 24: Rarefaction curves of *T. esculentum* microbial community based on 16S rRNA gene sequences generated using Roche /454 pyrosequencing (A) family and (B) genus at 0.03 distances.

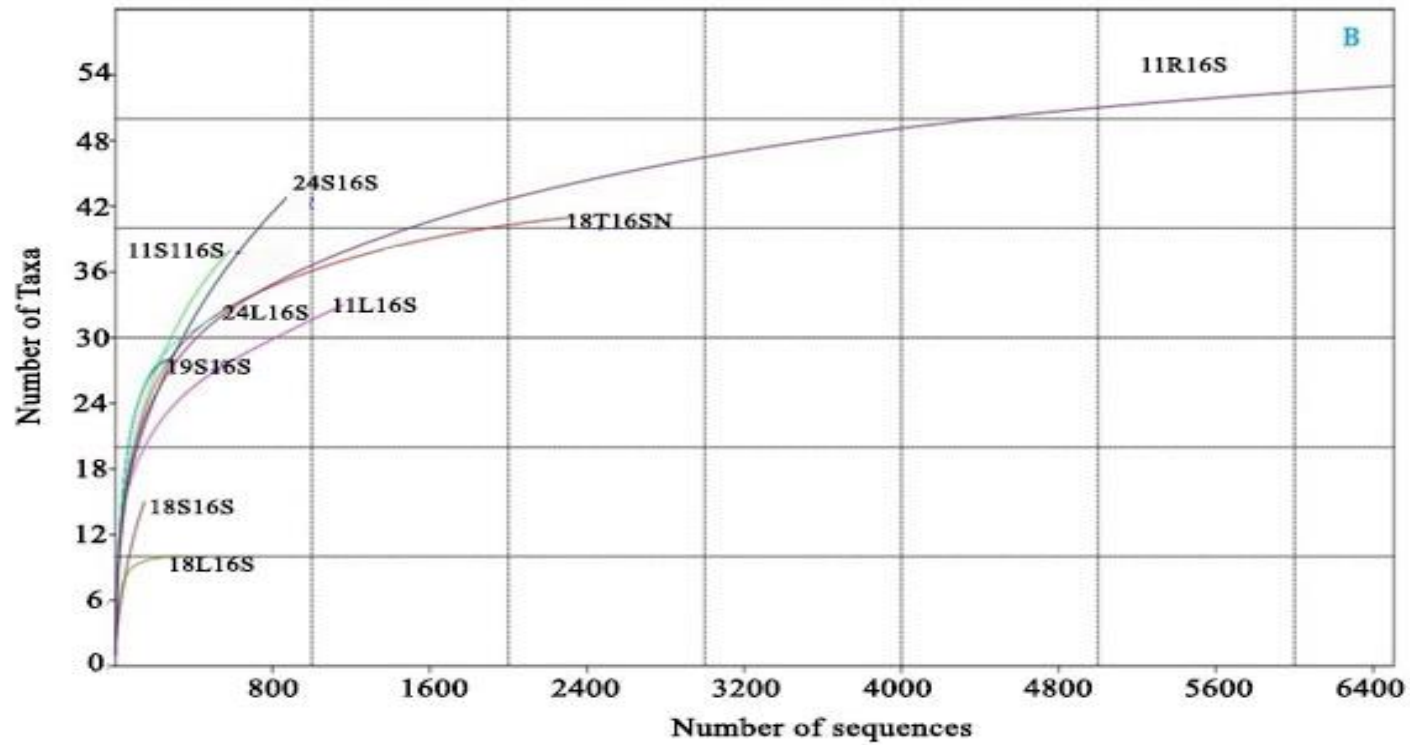


Figure 24 (Cont'd): Rarefaction curves of *T.esculentum* microbial community based on 16S rRNA gene sequences generated using Roche/454 pyrosequencing (A) Family and (B) genus at a 0.03 distance.

Rarefaction curves of the number of sequences of the samples collected in Omitara were compared. Rarefaction curves were generated via PAST software (Hammer, Harper, & Ryan, 2001) using a 95% identity cutoff for bacterial samples. The rarefaction was performed at the family and genus levels (Figure 24A and 24B), and rarefaction curves demonstrated that diversity differed between individual samples. Moreover, unexpectedly, endophytic microbial communities were more diverse in some stems and leaves than in the tubers.

Table 13: The number of OTUs and diversity indexes for total bacteria in all samples at three phylogenetic distances based on 16S rRNA gene sequences generated using Roche/454 pyrosequencing.

sampleID	distance	Number of clusters	clusters	Chao index	Shannon's diversity index (H')
11R16S	0.01		417	953.00	4.30
	0.03	1423	232	432.00	3.60
	0.05		153	232.00	3.02
11S16S	0.01		78	152.00	4.09
	0.03	145	61	99.00	3.78
	0.05		50	84.00	3.40
11L16S	0.01		132	263.00	4.60
	0.03	248	83	198.00	3.50
	0.05		72	192.00	3.20
18T16S	0.01		198	546.00	4.10
	0.03	528	118	225.00	3.40
	0.05		79	123.00	2.80
18S16S	0.01		36	72.00	2.90
	0.03	100	20	33.00	2.03
	0.05		14	17.00	1.34
18L16S	0.01		76	260.00	3.50
	0.03	183	46	67.00	2.90
	0.05		34	58.00	2.20
19L16S	0.01		45	104.00	3.60
	0.03	67	35	81.00	3.20
	0.05		32	71.00	3.00
19S16S	0.01		45	104.00	3.60
	0.03	67	35	81.00	3.17
	0.05		32	71.00	3.00
24L16S	0.01		78	180.00	4.00
	0.03	140	54	92.00	3.50
	0.05		42	63.40	3.08
24S16S	0.01		108	207.00	4.30
	0.03	228	79	137.20	3.60
	0.05		73	114.20	3.50

Tuberous roots (below ground) (18T16S and 11R16S) had higher species richness (953 and 546) and the highest Shannon's diversity index (4.10 and 4.3) respectively. This differentiated them significantly from the above ground (leaves and stems) (Figures 24A and 24B). Their (18T16S and 11R16S samples) asymptotes were starting to level off indicating that the majority of the *T. esculentum* microbial communities (species and families) have been reasonably characterized in our sampling efforts. Apparent disparity in differences for rarefaction curves for microbial communities originating leaves and stem (above ground) were observed. Their Shannon's diversity index, at 0.03 distance ranged between 3.7 and 2.03. This was indicated by the fact that many of the plots were not close to becoming asymptotic (Figures 24A and 24B). The rarefaction curves predicted that additional sampling will lead to significantly increased estimates of total diversity.

Table 13: Jaccard similarity matrix based on presence –absence data used as qualitative measurement for beta diversity

	18T16S	19L16S	19S16S	24L16S	11R16S	11S16S	18L16S	18S16S	11L16S	24S16S
18T16S										
19L16S	0.061538									
19S16S	0.061538									
24L16S	0.608696	0.134615	0.134615							
11R16S	0.727273	0.066667	0.066667	0.517857						
11S16S	0.612245	0.142857	0.142857	0.725	0.578947					
18L16S	0.177778	0.117647	0.117647	0.194444	0.163636	0.230769				
18S16S	0.23913	0.105263	0.105263	0.27027	0.259259	0.268293	0.444444			
11L16S	0.057143	0.84375	0.84375	0.122807	0.0625	0.131148	0.102564	0.093023		
24S16S	0.076923	0.642857	0.642857	0.138462	0.079545	0.144928	0.081633	0.096154	0.574468	

CHAPER 5: DISCUSSION

This study was carried out in the Kalahari Desert. It described and compared, *T. esculentum*'s endophytic bacterial community structures associated with leaves, stems and tuberous roots. Culture-dependent and culture-independent approaches were used to determine the density, the diversity and the distribution of these endophytic bacteria. The endophytic communities harboured in these different plant compartments showed different levels of diversity that were grouped in below and above ground communities. We have used molecular methods to analyse the total bacterial communities (16S rRNA gene) as well as the diazotrophic communities (*nifH* gene). The study provided insights on the endophytic microbial community associated with a Kalahari plant.

5.1 Numerical analysis of culturable bacterial communities living in *T.esculentum*

During their evolution, plants have developed expanded tools to screen microbial presence and to control their infection, therefore only specific microbes, so called “endophytes,” are able to colonise the internal tissues with insignificant or no host damage (Zgadzaj et al., 2015). Their population densities are highly variable, depending mainly on the microbial species and host genotype, developmental stage and environment conditions.

Within this study, densities of endophytic bacteria residing in root, stem and tuber tissues of *T. esculentum* were determined. The lowest microbial density was isolated in

stems collected in Harnas $(1.38 \pm 0.13)10^4 \text{ g}^{-1}$ fresh stem weight in 2013 while the highest was $(5.09 \pm 0.13)10^6 \text{ g}^{-1}$ fresh tuber weight isolated in Omitara in 2012 in a tuber. According to Mercado-Blanco & Lugtenberg (2014), the population density of endophytes was higher in roots than in any other plant organs. It has been estimated to the average of 10^5 CFU per g fresh weight in roots, whereas average values of 10^4 and 10^3 were reported for stems and leaf, respectively.

Our study revealed that the numerical average of endophytes associated with *T. esculentum* was higher than the average figure found in other studies (Mercado-Blanco & Lugtenberg, 2014). In stems, all obtained values were higher than average. This remark applied to tuber where the highest values were recorded. All were in the range of 10^6 . Additionally, plate counts were constantly significantly higher in *T. esculentum* tubers as compared to leaves and stems.

Below ground, root exudates are known to control rhizosphere microbial communities (Shia et al., 2015). However, there has not been any report of differences in *T. esculentum* fluctuation of root exudates with regard to climatic variations. In this specific case, it was not clear whether roots exudates were responsible for *T. esculentum*'s selective interior higher microbial density. Several ways by which microbes enter the plant interior have been described. However, in some instances plant phenotypic characteristics or even environmental conditions might favor bacterial establishment in the plant's interior tissues (Ramírez-Puebla et al., 2013).

As a result, a mutual plant-microbes nutrient exchange takes place between microbes and their host (Hardoim et al., 2011; Sessitsch et al., 2012) *de facto* increasing

their density. This was in agreement with our results where microbial density was higher in tubers compared to the above ground parts. The higher bacteria density in below ground tissues was probably due to a rich nutrient environment. It was therefore possible that these annual density differences might be the result of *T. esculentum* everchanging root exudates or its other phenotypic specificity. This should await further experimental verifications (Ramírez-Puebla et al., 2013).

Plant mature leaf tissues are subjected to temperature, humidity and UV radiation fluctuations. It can be predicted that these factors acted as a selective and variable components to the carbohydrate-rich leaf niche for the formation of an adapted bacterial endophytic communities (Ferrando, Fernández Mañay, & Fernández Scavino, 2012). Hence, the leaf dwelling microbial community should be able to acquaint to these physiological changes as *T. esculentum* evolves in its life cycle.

As in any other plant, while maturing, the accumulated carbohydrates in leaves are channeled into the grains (Yoshida, 1981). As a consequence, the mature leaf nutrient quality declines. Thus, the source of carbon becomes scarce for bacterial communities. Similarly, the low microbial density that occurred in the stems could be related to the key specific tissue, the phloem, that could provide a homeostatic sucrose rich environment and yet be poor in secondary metabolites (Ferrando et al., 2012; Emiliani et al., 2014). Accordingly, an ecological niche differentiation would be stimulated (Emiliani et al., 2014). These conditions may select for bacterial communities that would have adapted and developed a strong interaction with the *T. esculentum* communities in above ground niche.

Additionally, this interaction may be greatly influenced by the particular physiology of the plant host, *T. esculentum*. This would explain the lack of significant difference microbial densities between leaves and stems that was termed the above ground niche. The presence of endophytic bacteria was analysed as a function of the host plant organs. Our results indicated that certain plant parts prefer certain endophytes. However, these observations need to be confirmed, given that rarefaction curves for most host species indicated insufficient capture of microbial diversity (Figure 24).

5.2 Structure of culturable *T. esculentum* endophytic communities

Using culture-dependent methods, this study revealed that there were significant differences between *T. esculentum* microbial communities (below and above ground) composition ($p < 0.05$, Table 6). The results established that culturable detected taxa belonged to three bacterial phyla: *Proteobacteria*, *Firmicutes* and *Actinobacteria*. The wide distribution of genera accounted in *T. esculentum* has been documented as endophytes with beneficial effects in other plants. Most of them were involved in plant growth promotion (Ferrando et al., 2012; Grönemeyer, Burbano, Hurek, & Reinhold-Hurek, 2011; Emiliani et al., 2014).

Most sequences obtained in this study, have shown similarities with sequences occurring in environmental samples from arid, semi arid, cold desert and contaminated soils (Chowdhury, Schmid, Hartmann, & Tripathi, 2009). Similarities exist between these environments with respect to microbes' adaptability to water stress (Chowdhury, Schmid,

Hartmann, & Tripathi, 2009). It appeared that the belowground part (tuberous root) harboured diverse bacteria representing a broad phylogenetic spectrum compared to the aboveground part. Due to nutrient availability, the below ground part offered a more hospitable environment for endophytic microbial communities' survival.

Proteobacteria (*Alpha*-, *Beta*- and *Gammaproteobacteria*) was the most dominant group followed by *Firmicutes* and *Actinobacteria* to a lesser extent. A comparison with a similar study was conducted on *Catalystegia soldanella* and *Elymus mollis*, two sand dune species occurring in the Tae-An coastal Desert (Chungnam Province, South Korea) (Kim, 2005). The *Gammaproteobacteria* dominance was observed in the root and other parts of these plants. The microbial community diversity was greater than what was observed in *T. esculentum*. Six major phyla of the microbial domain were recorded. Though belonging to the same *Enterobacteriaceae* family, unlike *Enterobacter* in this study, the genus *Pseudomonas* made up the majority of the taxa.

In other studies on a poplar trees (Ulrich, Ulrich, & Ewald, 2008; Taghavi et al., 2010) and seagrass (Jensen, Kuhl, Glud, Jorgensen, & Prieme, 2005) growing in marginal soils, similar results were observed. The *Gammaproteobacteria* was the dominant class. However, an extra phylum was identified: the *Bacterioidetes* (Jensen, Kuhl, Glud, Jorgensen, & Prieme, 2005; Ulrich, Ulrich & Ewald, 2008; Ulrich, Ulrich, & Ewald, 2008; Taghavi et al., 2010). Groenmeyer, et al. (2011) conducted a study on root endophytes occurring in some of the agricultural crops (pearl millet, sorghum and maize) in Kavango East (Namibia). This region is known for its prolonged dry seasons. Similar results (phyla)

were obtained. However, the majority of the isolates belonged to *Firmicutes* and *Actinobacteria* while *Proteobacteria* were poorly represented. These differences can be attributed to differences in sampling periods. *Firmicutes* and *Actinobacteria* are known soil dwellers and spore formers (a surviving strategy under arid conditions). The microbial endophytic communities' diversity in plants might differ between plant species and locations. Conversely, *T.esculentum* bacterial endophytic communities shared several similar adaptation traits, in term of phyla composition, with endophytic communities of other plant species.

There are a number of factors that are known to influence the plant microbial endophytic population such as the soil condition and the number of phytopathogens (El-Deeb, Fayez, & Gherbawy, 2012). In Desert habitats, plant interactions with microbes appear to be of uttermost importance in obtaining inorganic nutrients or growth-influencing substances (El-Deeb et al., 2012). Generally, plants growing in unique environmental settings or interesting endemic locations possess novel endophytic microorganisms which can supply new leads (El-Deeb et al., 2012). Hence, it can be suggested that members of these endophytic phyla described above that were constantly found in plant specific organs across host plant species, have a critical function in microbial communities (Emiliani, et al., 2014).

According to Fierer et al. (2012) and Makhalanyane et al. (2013) *Alpha*-, *Beta*- and *Gammaproteobacteria* are characteristic of soils with higher rates of organic carbon inputs. Acosta-Martinez, Dowd, Sun, & Allen (2008) suggest *Gammaproteobacteria* are

known to frequently associate with C and N cycling in soils. This coincided with the rainy season and the subsequent littering of plants when *T.esculentum* re-sprouting occurred and the desert was greening.

5.2.1 *T. esculentum* above ground culturable microbial community diversity

The above ground endophytic bacterial community was represented by all the three phyla (*Proteobacteria*, *Firmicutes* and *Actinobacteria*) comprising 38% of the total *T. esculentum* microbial community (Table 6). Almost 42% of all isolated taxa recorded in *T. esculentum* were confined in leaves (*Lysinibacillus*, *Bacillus*, *Streptomyces*, *Acinetobacter*, *Enterobacter*, *Klebsiella*, *Kosakonia*, *Pantoea*, *Stenotrophomonas*, and *Pseudomonas*). A lower diversity was found in stems where several taxa that were recorded in leaves were absent in stems. It was found that members of these three phyla were the most abundant in the endophytic communities of wildflower (*Crocus albiflorus*) (Reiter & Sessitsch, 2006); radish leaf (Seo, Lim, Kim, Yun, & Cho, 2010); *Arabidopsis* and Citrus leaves (Bodenhausen et al., 2013); in xerophilous moss (*Grimmia montana*) (Liu et al., 2014); olive oil leaf (Muller et al., 2015). *Proteobacteria* seem to be highly abundant in the leaf endophytic communities analyzed. *Gammaproteobacteria* was the most abundant class (98%) of *Proteobacteria* (Table 6) in the endophytic community of *T. esculentum* above ground. A very low or even an absence of α and β *proteobacteria* is worth noticing (Table 6). It showed a pattern of separation of the classes of the phylum *Proteobacteria* between above and below ground plant parts during plant colonization.

Similarly, microbial diversity patterns within samples held the same pattern. Besides the highly abundant *Proteobacteria*, the aboveground endophytic community also included a small proportion of *Actinobacteria* (2%) and *Firmicutes* (10.7%) (Table 6). These minor phyla were also reported as minor components of the endophytic communities of citrus leaves (Sagaram et al., 2009) and *Thlaspi goesingense* shoots (Idris, Bochow, Ross, & Boriss, 2004).

In desert soils, the class *Alphaproteobacteria* has been found to be the most prevalent (Makhalanyane et al., 2013). In this study, it was only prominent in the tuber and not in the leaves. Similar findings were observed by Emiliani, et al. (2014) while studying bacterial communities associated with *Lavandula angustifolia* and their impact on essential oil production. They suggested this component of the microbial community might have diffused from the seeds through to the plant leaves during the plant development.

While investigating the bacterial communities associated with leaves and roots in *Arabidopsis thaliana*, it was hypothesized that bacteria first colonize the leaf surface and then selectively diffused through stomata (Bodenhausen et al., 2013). This might most likely have taken place in *T. esculentum* since it is a prostrate plant. It is possible, on the other hand, for bacteria to migrate from tubers to leaves. It has been shown that GFP-tagged beneficial bacteria such *Rhizobia* inoculated in the soil were found in leaves (Bodenhausen et al., 2013). These specific organ microbial communities have to be adapted to the specific plant environment they colonize. Therefore, it has been suggested

that metabolic potential are most likely to be adjusted to the new environment compared to microbial soil dwelling counterpart (Mitter et al., 2013). However, obligate endophytes are endowed with a less rich metabolic arsenal compared to facultative endophytes. The latter are equipped with a richer arsenal for they are involved in competition with the rhizospheric as well as endophytic plant colonizers (Mitter et al., 2013).

In the present study, plant-specific microbial distribution might attest that *T. esculentum* microbial community colonization was not passive, at least at a certain developmental stage. Studies have proposed a two stage model in plant endosymbiont colonization. The first microbial intake involves a microbial community determined by soil abiotic factors in the plant vicinity. This microbial intake is then modified by plant root secretions that promote organotrophic bacteria. The second stage allows host genotype-specific genetic microbes (Bulgarelli et al., 2013) and (Emiliani et al., 2014). This assumption was later confirmed by Edwards et al. (2015) in their study on the structure, variation and assembly of the root-associated microbiomes of rice.

5. 2.2 Below ground culturable *T. esculentum* microbial community

In our study, it can be assumed that due to the large number of cattle in Omaheke region as evident from the obvious omnipresence of cow dung, a lot of enteric bacteria genera can be expected in the soil. Kennedy, Choudhury, & Kecskés (2004) postulated that plants may have developed a flora well adapted to an animal–soil–plant rhizosphere nutritional

cycle. These enteric genera included *Klebsiella*, *Enterobacter*, *Citrobacter*, *Pseudomonas* and several others yet to be unidentified (Kennedy, Choudhury, & Kecskes, 2004).

The smooth plant–*Enterobacteriaceae* interaction was facilitated by the fact that *Enterobacteriaceae* are renowned for their motility which is an important factor in plant colonization. Another attribute associated with the *Enterobacteriaceae* family members is the cellulase and pectinase enzyme production. They act as virulence factor for bacteria pathogens and they are believed to be involved in host plant penetration by Plant Growth Promoting Bacteria. *Enterobacteriaceae* have the capacity to metabolize a widespread variety of sugars. The possession of these characteristics may have contributed to the ability of these isolates in colonizing roots and subsequently other parts of the plant.

Firmicutes comprise Gram positive and low GC content bacteria. They are commonly isolated from plant tissues. They were the second largest phyla that accounted for 23.7% of the total isolates in *T. esculentum*. Of the three classes, *Bacilli*, *Clostridia* and *Erisipelotrichi* (Bueche et al., 2013) only the first two spore formers were reported in *T. esculentum*. Statistical analyses have shown a significant difference between above and below ground microbial composition (Table 6).

Member of the *Firmicutes* phyla have been reported in desert soils and plants mainly the *Clostridia* class (Makhalanyane et al., 2013). This class has shown to be dominant in rhizospheric soils associated with Antarctic vascular plants (Makhalanyane et al., 2013). This emphasized *Firmicutes* importance in arid environments (Makhalanyane et al., 2013). In *T. esculentum*, *Clostridiales* were poorly accounted for;

the genus *Bacillus* was the most dominant.

Members of the species of *Bacillus* are useful and widespread in agricultural fields. They directly or indirectly contribute to plant productivity (Kumar, Prakash, & Johri, 2011). *Bacillus* dominates plant microbiomes under arid conditions such as in Egypt where *Pseudomonas* cannot survive (Köberl et al., 2013). The genus *Bacillus* is endowed with particular physiological traits that contribute to its survival and ubiquitous character such as multilayered cell wall, stress resistant endospore formation, and secretion of peptide antibiotics, peptide signal molecules, and extra cellular enzymes (Kumar et al., 2011). According to Malfanova et al. (2011), accumulating evidence suggest that bacteria belonging to the genera *Bacillus* have been shown to be able to produce siderophores , IAA and antibiotic compounds which can be effective against deleterious effects of phytopathogen agents (Kumar et al., 2011).

Recently, it was reported how *Bacillus* and *Paenibacillus* can effectively suppress pathogen under arid conditions (Köberl et al., 2013). Moreover, it was established that more than 8.5% of the plant root-colonizing *B. amyloliquefaciens* FZB42 genome is committed to antibiotics and siderophores synthesis by pathways not involving ribosomes (Chen et al., 2010). It has been recently reported that bacterial antibiotics and lipopeptides are responsible for regulators and support biofilm formation, signaling, motility, and acquisition of micronutrients at sub-inhibitory concentrations (Berg et al., 2014). Several studies have attested that plant genotypes can affect the accumulation of microorganisms that help the plant to defend itself against pathogen attacks (Berendsen, Pieterse, &

Bakker, 2012).

Finally, Oa et al. (2008) also demonstrated that the excretion of malic acid via plant exudates was responsible for *Bacillus* recruitment. Considering *Bacillus* dominance in *T. esculentum* microbial community, it can be hypothesized that *T. esculentum* might be selectively producing malic acid to attract and harbor a highly *Bacillus* dominated community.

Moreover, it has been established that some members of the genus *Bacillus* can produce the α -amylase enzyme responsible for starch hydrolysis (Puspasari, Talib, Daud, & Tasirin, 2014). Conversely, it would be premature to correlate the strong presence of *Bacillus* in *T. esculentum* tuber with the starch hydrolysis trait of some *Bacillus* genera. There is no data to support *Bacillus*' presence with the amount of starch present in the tuber. The relative abundance of the genus *Bacillus* in *T. esculentum* and particularly in the tuber could be due to the plant tolerance of this specific genus and its recruiting abilities for specific functions.

Actinobacteria were not highly represented but they were mostly found in tubers (2.14% of the total microbial population recorded in this study) with three genera namely *Arthrobacter*, *Curtobacterium*, *Streptomyces* and unclassified *Actinomycetales*. Their low number might be partly due to the non-selective *Actinobacteria* media used. *Actinobacteria* (*Actinomycetes*) are known for their production of various active natural compounds such as anti fungal and bactericidal compounds or even plant growth promoting substances (Jiang et al., 2013). Endophytic *Actinobacteria* isolated from

wheat were able to suppress fungal pathogens both in *vitro* and in *vivo*. They induced systemic acquired resistance (SAR) in *Arabidopsis thaliana* (Conn, Walker, & Franco, 2008b).

Endophytic *Arthrobacter* have been previously found in wheat (Conn, Walker, & Franco, 2008a), root and stems of Black pepper (*Piper nigrum* L.) (Aravind, Kumar, Eapen, & Ramana, 2009). More than 70 species with agricultural, medical, and environmental functions have been so far recognized by researchers (Fu et al., 2014). Endophytic *Streptomyces* strains with the ability to produce IAA in *vitro* have been isolated from plants occurring in poor sandy soil and arid climatic conditions of the Algerian Sahara (Gougjal et al., 2013).

As the major decomposer of chitin, polymeric carbon and other organic detritus, *Actinobacteria* probably originated from cattle faeces or soil, which invades the host plants from the phyllosphere. Their low numbers can be explained by the fact that *Actinobacteria* are subjected to dominance of the *Gammaproteobacteria* and *Firmicutes* in oligotrophic environments.

These phyla that were constantly found in *T. esculentum* specific organs across samples can be predicted to have a critical function in microbial communities. The dominance of *Proteobacteria* and *Firmicutes* reflect the ability to survive desert conditions that are known for their nitrogen and carbon deficiency. Hence, describing the interaction of these bacteria with the host plant would be of great interest with regards to our knowledge about the plant–microbe interactions and thereafter their

economic potential and environmental benefits in agriculture. Very little is known about *T. esculentum* interactions with the microbial community in natural ecosystems.

5. 2.3 *T. esculentum* culturable diazotrophic bacteria

Using microbiological and molecular methods, the existence of putative nitrogen fixing microbial communities in *T. esculentum* has been confirmed. Most of these genera were already known to include species containing *nifH* sequences, such as *Bukolderia* (Chen et al., 2005) that was found to nodulate Mimosa in South America, *Xanthobacter* (Malik & Claus, 1979), *Azospirillum* (Steenhoudt & Vanderleyden, 2000), *Stenotrophomonas* (Ramos, Yates, Pedrosa, & Souza, 2011), *Bacillus* (Ding, Wang, Liu, & Chen, 2005), *Ideonella* (Noar & Buckley, 2009).

In previous studies, endophytic diazotrophs were mostly dominated by the Phylum *Proteobacteria* (Kizilovaa, Titova, Kravchenko, & Iutinskaya, 2012). This was in agreement with our results. All viable strains in our study showed an indication of nitrogen fixation. Regardless of their plant part origin, they were all positive for *nifH* amplification. Therefore, the same culturable distribution along the two parts (below and above ground) was applied to the *nifH* distribution. The presence of nitrogen fixing trait in all endophytes would point to a yet unexplored type of symbiosis with *T. esculentum*, as an adaptation to the poor-nutrient and harsh Kalahari environment.

Dakora, Lawlor, & Sibuga, (1999) conducted a ¹⁵N tracer experiment to

investigate the site of nitrogen fixation in *T. esculentum*. The fixed ^{15}N was never detected in the plant internal tissues. They could not decipher the enigmatic origin of the high protein concentration in *T. esculentum*. To promote yield and plant growth, diazotrophic endophytes nitrogen transport to the host plant is important in addition to the occurrence of high nitrogen fixation activity (Ohya et al., 2014). In *T. esculentum*, this mechanism has never been elucidated. Dakora and his colleagues (1999) only considered the soil as the only source of diazotrophic bacteria. They have underestimated the role of seed associated microorganisms.

Seeds are the plant continuation organ and can persist for years (Liu, Zuo, Xu, Zou, & Song, 2012) and (Truyens, Weyens, Cuypers, & Vangronsveld, 2015). They harbour a variety of beneficial bacteria (Liu et al., 2012) capable of playing a role in seed preservation and preparation of the environment for germination (Truyens et al., 2015). They are transferred from parents to the progeny to benefit the next generation (Truyens et al., 2015). Seed and seedling endophytic bacteria diversity have been extensively studied in *Eucalyptus* (Ferreira et al., 2008), maize (Liu et al., 2012), rice (Kaga et al., 2012). In barley, seeds diazotrophs were isolated (Zawonzik, Vazquez, Herrera, & Groppa, 2014). Most bacteria accounted in seeds belong the *Proteobacteria* (80 genera). *Firmicutes* and *Actinobacteria* were well represented as well. It has not yet been established whether *T. esculentum* endophytes were the results of vertical transmission. Since Dakora and his colleagues (1999) failed to prove any nitrogen fixation site, we can hypothesise that nitrogen fixing bacteria associated with *T. esculentum* might be vertically transmitted.

Comparative analysis of the *nifH* gene sequences obtained from this study with and their closest match in NCBI database could provide imperative information concerning the phylogenetic position of nitrogen fixing bacteria (Kizilovaa et al., 2012). Phylogenetic tree topology incongruency between *nifH* and 16SrRNA were detected. The *nifH* gene did not always align with its 16S rRNA matches. This was also previously reported by Minerdi, Fani, & Gallo (2001) in *Burkholderia* species. Ecological isolation coupled with host and their symbionts intimate physiological interactions can generate divergences and bacterial diversity in special niches (Barraclough, Hughes, Ashford-Hodges, & Fujisawa, 2009).

In this study, one would argue that plant occurring in extreme environments have evolved and developed mechanisms to highly regulate their microbial interaction. As a result, a mutual reciprocal host-bacteria recognition and coexistence have been established (Zgadzaj et al., 2015). In the course of evolution, they may have developed a genetic rearrangement that might have taken place occasioning a more complicated evolutionary history (Laguerre et al., 2001). This is an important ecological adaptation that confers to the plant a selective advantage to acquire bacterial strains with these particular vital traits. At the origin, it can be proposed that a lateral gene transfer might have been responsible for this phenomenon (Doty et al., 2009). Another alternative explanation is that these bacteria might have originated from a common bacterial ancestor, and were subject to an evolutionary pressure with regard to nitrogen acquisition functional traits (Dedysh, Ricke, & Liesack, 2004). This genetic rearrangement might have been at the origin of diversification and in structuring *T.*

esculentum natural diazotrophic populations.

Few studies have addressed the direct correlation between the nitrogenase activity and diazotrophic diversity. Contradictory reports exist. Deslippe & Egger (2006) reported a poor relationship between *nifH* gene diversity and nitrogenase activity. Burgmann, Widmer, Sigler, & Zeyer (2004) reported completely the opposite findings. Later, in his studies, Canton et al (2012) concluded that nitrogen fixation was driven by the nitrogen loading in the environment. His analysis of acetylene reduction suggested that nitrogen fixation activity was higher in poor nutrient environments. Furthermore, in pristine environments, diazotroph communities showed a higher nitrogen fixing activity level. These conclusions might partly explain *T. esculentum*'s rich protein seeds.

Additionally, previous studies have verified that in sediments, an NH_4^+ elevated concentration would be responsible for a lower nitrogenase activity by minimizing energy consumption. In case of NH_4^+ removal, the diazotrophic bacteria activity is increased (Burns, Zehr, & Capone, 2002; Dixon & Kahn, 2004). Equally, it has been reported that regardless of the amount of nutrient in some diazotroph microbial communities (Burns, Zehr, & Capone, 2002; Dixon & Kahn, 2004), some diazotroph bacteria would develop the ability to independently perform their nitrogenase activity (Burns, Zehr, & Capone, 2002; Dixon & Kahn, 2004). Moreover, it was substantiated that in case of highly diverse diazotroph population and highly specialized niches like in *T. esculentum*'s case, some predominant diazotroph might be inhibited by the the presence of nitrogen in their surroundings giving an opportunity to other to efficiently

and independently increase their nitrogenase activity (Bagwell, Dantzler, Bergholz, & Lovell, 2000). Results from this study strongly hypothesize that during its life time, *T. esculentum*'s diazotroph microbial community would remain unchanged but would rather have intermittent diazotroph activity. However, methods used in this study could not allow direct comparison of nitrogenase activity and active diversity.

Furthermore, Gaby & Buckley (2014) established that the genetic divergence of *nifH* and 16S rRNA genes did not correlate well at sequence dissimilarity values used commonly to define microbial species. They concluded that in some instances strains with <3% sequence dissimilarity in their 16S rRNA genes could have up to 23% dissimilarity in *nifH*. This assumption would explain this higher *nifH* gene incongruence in our study. It was probably due the high disparities in 16S rRNA sequences dissimilarities.

5.3 Microbial community structure

5.3.1 Cluster analysis

The uniqueness of root endophyte communities was illustrated in the hierarchical clustering-based analysis of the dominant genus from each sample (Fig. 17). Clustering bacterial composition between the three plant compartments (leaves, stems and tuberous root) originating from 3 different locations displayed a striking separation between the above ground (leaves and stems) and below ground (tuberous roots) parts. These results

are in line with results found elsewhere.

Pisarska & Pietr (2015) in their study on endophytic bacteria inhabiting the same maize cultivar grown in different location concluded that plants were colonized with different associations of endophytic bacteria. They suggested the plant genotype regulates endophytic bacteria from local agro-environments. The plant microbiome is dynamic, selective and controlled by various factors such as the rhizosphere pH and the presence of chemical signals from bacteria (Lakshmanan, Selvaraj, & Bais, 2014). The distribution of endophytes is assumed by the plant. According to Mandl et al. (2014) vines take up yeast from soil and channel them to stems and skins of grapes. It can be suggested that bacterial communities associated with roots are not random samples from a common pool of genera, but that these two microbiota (above and below ground) differed in their microbial composition. This can be attributed to *Enterobacteriaceae* and *Bacillaceae* high abundance in the tuber. These families were highly dominated by the genus *Acinetobacter*, *Enterobacter* and *Stenotrophomonas*, *Pseudomonas* and with a big number of Unclassified *Enterobacteriaceae*. PCA also confirmed the clustering results that the two groups of samples, from below ground and above ground were clearly different and indistinguishable.

5.3.2 Principal Component Analysis of microbial communities analysis

The Principal Component Analysis (PCA) consist of a multi-dimensional coordinates

analysis capable of taking data from large amounts of observations and condensing into a reduced amount of components that would allow to determine which taxonomic group vary with one another. Thus, assumptions can be made about the relationships of particular populations. Hence, it would be possible to evaluate which of the original observations solidly contribute to those components and possibly begin the assemblage of functional associations (Raymon et al., 2013). Principal component analysis was performed using PAST software (Hammer, Harper, & Ryan, 2001) based on covariance of the 16S rRNA data (relative abundance of all recorded genus). PCA on *T. esculentum* microbial community analysis revealed that all the root samples clustered together and leaves and stems were clustering the closest. As expected, PCA analysis failed to separate these 2 groups, suggesting that the overall bacterial community structure and composition between the 2 groups were indistinguishable (Figure 18). At the onset of this study, we expected some differences between Harnas and Omitara located in the commercial farmer area and Otjinene which is a communal farming area. Despite these differences, bacterial communities clustered together as per their host tissue regardless of their location.

Considering the composition of the first component (PC1), the taxa present seemed to be fairly representative of the plant growth promoter in the most represented phyla *Firmicutes* and *Proteobacteria*. It can be suggested that *Bacillus*, *Acinetobacter* and unclassified *Enterobacteriaceae* were selected for their plants growth promoting traits to compensate for the lack of nutrients in the Kalahari soils (Wang, D'Odorico, Ringrose, Coetzee, & Macko, 2007). These microbes are adapted to elastic environments as their cycle can survive between wet and dry season.

The microbial dominance in the second component differ from the PC1's. In PC2, we had *Klebsiella* which was phylogenetically related to *Acinetobacter* and *Enterobacter*. This strong differentiation of endophytic communities belonging to the *Enterobacteriaceae* family suggested that their selection is strongly influenced by the recruitment of the rhizosphere microbiome and the genotype of the host (*T. esculentum*).

5.3.3 Diversity indices

Recently, it has been shown that when epiphytic bacteria are isolated from the host and compared among different species, host phylogeny is more related to the composition of the epiphytic community than to the region of origin (Hollants, Leliaert, Verbruggen, Willems, & De Clerck, 2013). It appeared also true when bacteria community composition changes seasonally (Figure 8 (A) (B) (C)) (Aires, Serrão, Kendrick, Duarte, & Arnaud-Haond, 2013; Stratil, Neulinger, Knecht, Friedrichs, & Wahl, 2013). It appeared to be the case with *T. esculentum* whereby, according to PCA, though clustering together, the microbial composition was never the same on a yearly basis. It can be predicted that *T. esculentum* is a reservoir of bacterial diversity.

In our study, the change in Shannon–Wiener index ranged from 1.22 in Otjinene in 2013 to 2.69 in tubers in Omitara in 2011 (Figure 5 and Table 8). After one way ANOVA analysis and Tukey test, it was concluded that there is a significant difference between the tuber and leaves ($p=0.005$) and stems ($p=0.006$) microbial communities (Table 8). This trend reflects that bacterial diversity became abundant presumably because

of the combined effects of root exudates and environmental conditions especially pedoclimatic conditions (Marasco, et al., 2013). The reduced microbial diversity in the above ground plant tissues may reflect specific physiological requirements to enter the interior of the leaf and stem to establish endophytic populations.

In the tuber, the higher microbial diversity might be the result of the rhizosphere effect. It has been demonstrated that it is pronounced especially in nutrient-poor soils and under severe abiotic stresses, as previously observed for herbaceous and arboreal plants grown in arid lands (Marasco et al., 2012; Marasco et al., 2013). This might be the case in the Kalahari Desert known for its poor sandy soils with very little water retention capacity coupled with its extreme environmental conditions.

The higher values of the Simpson dominance in the tubers (0.84 in OM2013, 0.87 in OM2012, 0.85 in HARN2014, 0.90 in HARN2011 and 0.80 in OTJ2011) could be explained by the tuber's species specificity with genus such as *Streptococcus*, *Lactococcus*, *Curtobacterium*, *Ochrobactrum*, *Burkholderia*, *Rhizobium*, *Achromobacter*, *Azomonas*, *Consenzaea*, *Trabulshiella* and the unclassified *Pseudomonadaceae* that were at least recorded once. However, a comparison of the dominance values for microbial communities in different plant parts of *T. esculentum* using the Box-plot coupled (Figure 5) with ANOVA (Table 9), it was revealed there was no significant difference in microbial community dominance between leaves, stems and tubers ($p = 0.20$) (Table 9).

5.3.4 Phylogenetic analysis

Using the 16S rRNA, 49% the threshold percentage (98.7-99%) at 16S rRNA loci has shown to correspond to phylogenetically delimited species in the Bacteria domain. Thus showing that in this study; there might be a high proportion of putatively new species in endophytes associated with *T. esculentum*. Furthermore, there were no clustering by geographic origin or plant organs observed in our samples. This was in line with Stackebrandt and Ebers (2006) in Burbano, Grönemeyer, Hurek, & Reinhold-Hurek (2015) and Figueiredo et al. (2011) who considered this as consistent boundary in species delimitation. The reliability of the alignment was confirmed using the E-value that indicated that the probability of obtaining an alignment a great number of reported taxa (49%) showed the identity similarity values below 98.7-99% (Table 15).

5. 4 Bacterial communities associated with leaves, stems and tubers of *T. esculentum* determined by 454 pyrosequencing techniques

This study provides the first next-generation sequencing survey of the bacterial community in the different organs of *T. esculentum*. In the first part of this study, a culture dependent method was used. The pyrosequencing technique not only supported the previous study, but provided new insight into the bacterial communities of *T. esculentum*.

5.4.1 Bacterial community structure revealed by 16S r RNA gene amplicon sequencing.

Studies of plants endophytic bacteria have established the significance of plant microbial interaction and their subsequent adaptation to diverse ecosystems. However, little is known about the colonizing species, the relationship between these bacteria and wild plants, and the possible benefits of this interaction.

In this study, almost 63% of the sequenced members of the bacterial communities associated with *T.esculentum* were examined at phylum level on the basis of the RDP Bayesian Classifier (Cole, et al., 2014). This value was higher than the score obtained in the pyrosequencing study of spinach phylloepiphytes (54%) (Lopez-Velasco, Welbaum, Boyer, Mane, & Ponder, 2011) and lower than the results obtained in the pyrosequencing of the rhizosphere associated with the potato root (75%) (Manter, Delgado, Holm, & Stong, 2010). These differences were probably due to primers specificity. Avoiding chloroplast amplification would, for sure, increase the number of reads. In fact, the primer 799f known to avoid chloroplasts and cyanobacteria matches only 62% of the RDP sequences when the Probe Match Tool on RDP-II is used (Cole, et al., 2014). Another factor to be taken into account is the sequencing depth (Santhanam et al., 2014). Due to our limited sampling depth and lack of replicates, we were not able to determine whether differences in microbial communities in plant parts are merely due to our sampling methods or the primers used. However, our results have clearly shown that culture

independent techniques can reveal previously undetected microbial communities and highlight their putative functional importance. Plotting the rarefaction curve was done (Figure 24) to assess how complete the sampling was for the microbial community associated with *T. esculentum*. In some case we failed to approach an asymptote for bacteria sequences in the sample, at the family level.

5.4.2 *T. esculentum*-associated bacterial communities

A comparison of the communities associated with the leaves, stems and roots revealed both ubiquitous and organ specific bacteria groups. Across all samples, endophytic bacteria colonizing *T. esculentum*, 5 different phyla were detected and they were distributed as follow: *Firmicutes* (50.3%), *Proteobacteria* (38.32%), *Fusobacteriia* (5.7%), *Actinobacteria* (4.380%) and *Bacteroidetes* (1.3%). In the *Proteobacteria* phyla α , β , γ and ϵ *Proteobacteria* were identified. There were clearly some most prominent distinctive taxonomic groups in our data. Only two phyla, *Firmicutes* and *Proteobacteria* of the 5 detected comprised nearly 89% of the total bacterial assemblage. Equally, 4 families of the 18 recorded, namely *Enterobacteriaceae*, *Bacillaceae*, *Pasteurellaceae* and *Fusobacteriaceae*, constituted 91% of the total *T. esculentum* bacterial community. Surprisingly, more than 50% of the sequences were unclassified. Moreover, all the 5 phyla were detected in the aboveground plant parts, while only three (*Proteobacteria*, *Firmicutes* and *Actinobacteria*) were accounted in below ground plant parts.

Under conditions of mixed biotic and abiotic factors, physiological pathways in plants that determine an appropriate bacterial outcome of plant–microbe interactions are given a priority (Schenk et al., 2012). Therefore, it would not be surprising that there were high relative abundances of *Firmicutes* and *Proteobacteria* in *T. esclentum*. The two bacterial phyla were also known to colonize other plants (Jorquera, Shaharoona, Nadeem, Mora M, & Crowley, 2012) such as wheat (Andersona & Habigerb, 2013), *Arabidopsis* (Bodenhausen et al., 2013) and rice (Edwards et al., 2015). This led to the proposal that plants in synergy with multiple endophytes can better contribute to plant disease suppression (Schenk et al., 2012).

Some plant endophytes, in adverse conditions, can also hasten seedling germination, promote plant establishment (Saharan & Nehra, 2011) and improve plant growth (Vetrivelkalai, Sivakumar, & Jonathan, 2010). Species of *Proteobacteria*, *Firmicutes* and *Actinobacteria* are known to exert beneficial effects to their plant host (Edwards et al., 2015). It has been documented that members of the phyla *Firmicutes*, *Proteobacteria* and *Actinobacteria* were capable of suppressing plant diseases (Edwards et al., 2015).

Among all representative sequences, almost 50% of them showed 99–100% sequence similarity, whereas more than 50% at the genus level could not be associated with any genus in the RDP Database sequence similarity. According to Stackebbrandt and Ebers (2006) in Burbano et al. (2015), for bacterial identification a threshold of more than 97% sequence similarity should be acceptable. However, less than 97% similarity of the

16S rRNA gene with known bacterial species is justified for a novel bacterial species. The threshold for new genera varies between 93 and 95% among bacterial phyla. It can be anticipated that more than half of the bacteria sequences detected in *T. esculentum* could be new species for they are very far from fitting in these set thresholds.

5.4.3 *T. esculentum* above ground microbial community diversity

The above ground *T. esculentum* bacterial community is dominated by *Proteobacteria* at the phylum level (57%) followed by the *Fusobacteria* (23%), *Firmicutes* (12%), *Bacteroidetes* (5%) and *Actinobacteria* (3%). The number of phyla was much higher compared to the one obtained using the culture dependent method. This difference could be partly associated with injury from oxidative stress caused by the sampling and disruption of plant tissue methods during isolation and sampling (Muresu et al., 2012). An experiment was conducted using buffers supplemented with scavenging systems to prevent damage from reactive oxygen species (ROS). It yielded dramatic increase in culturable endophytes (Muresu et al., 2012). In addition, the above ground has a higher diversity at a phylum level while it has shown to have a lower diversity at the species level (Figure 27). The aboveground rarefaction curve samples (18L16S, 18S16S, 19S16S, 24L16S, 24S16S, 11S16S, and 11L16S) did not level off. It can be suggested that this sequencing depth was still not enough to cover the whole *T. esculentum* bacterial diversity. Hence, further sequencing would be valuable to detect more species.

A series of diversity indexes (Chao and Shannon diversity and E index) suggested that the bacterial communities above ground *T. esculentum* were shown to have an apparent high diversity (Table 11). At 1% cutoff, the Chao index indicated that, aboveground, the 18L and 16S (Chao index 259.66) and 11L and 16S (Chao index 262.88) samples have the richest communities (Table 11). The Shannon index also indicated that aboveground samples 11L16S and 24S16S were the most diverse, with a Shannon index of 4.60 and 4.30 respectively. In particular, the high bacterial diversity in sample 11S16S, 19S16S, 24S16S and 24L16S was noticeable by rarefaction curves. This is due to the presence only of *Pasteurellaceae*, *Fusobacteria* and *Bacteriodetes* only in the above ground plant parts.

Within the *Proteobacteria*, the *Gammaproteobacteria* endophytic community was the most dominant with *Haemophilus* being the most abundant genus (110 sequences). The genera *Pantoea* and *Haemophilus* were found in leaves and stems while *Enterobacter* is ubiquitous. *Alphaproteobacteria* (*Brevundimonas* and *Ochrobactrum*) were absent in the stems but present in leaves. The phyla *Epsilonbacteria* and *Fusobacteria* were also present in tuberous root and leave but absent in stems. However, this was not consistent across the samples. The bacterial community in leaves was more diverse than in stems. There was a larger number of different genera present in leaves and absent in stems. This is probably due to a selective pressure that the plant exerts on their associated bacterial population (Muresu et al., 2012). Endophytic bacteria of plant colonize specific ecological niches in different tissues, which might be the reason for the diversity observed within the bacterial community.

The most abundant bacteria detected above ground and predominantly in leaves are the *Pasteurellaceae* and *Fusobacteriaceae* (Figure 17). *Haemophilus* was the only member of the *Pasteurellaceae* family recorded in *T. esculentum*. The family *Pasteurellaceae* is the single constituent family of the order *Pasteurellales*. Species of this family represent a diverse group of mostly commensal and few pathogenic bacteria within the class *Gammaproteobacteria* (Naushad et al., 2015). Species of this family are well-known to specialize as nonpathogenic members of the resident flora of the healthy upper respiratory tract and oral cavity of birds and mammals (Di Bonaventura, Lee, DeSalle, & Planet, 2010; Naushad et al., 2015). The genera *Haemophilus* contain species responsible for human bacteremia, pneumonia, acute bacterial meningitis, and the sexually transmitted disease chancroid (Di Bonaventura et al., 2010). However, depending on the environment conditions, it has been established that many bacterial species are able to switch between different lifestyles (Rezzonico, Smits, & Duffy, 2012). Members of the *Pasteurellaceae* family were also isolated from the *Caragana microphylla*, a drought tolerant legume occurring in poor, well drained sandy soils in the desert grassland of the Ningxia (China) (Dai, Liu, & Wang, 2013). In the case of biofertilization manufacturing, bacteria that are potential human pathogens should be avoided.

In previous studies, the *Pasteurellaceae* family was detected in hyphae of fungal foliar endophytes where they were implicated in hyphal promoting growth and the establishment of ectomycorrhizal growth (Hoffman & Arnold, 2010). It was established that these ectosymbiotic bacteria had the ability to limit the pathogenicity of *Fusarium* to cause vascular wilt in lettuce (Hoffman & Arnold, 2010). Furthermore, the removal of

bacterial partners from fungal spores would suppress fungal growth and development (Hoffman & Arnold, 2010). Yet, their role in *T. esculentum* has to be elucidated whether associated with ectomycorrhizae or not.

Endophytic populations vary from plant to plant and from species to species. Within the same species it not only varies from region to region, but also differs with change in climatic conditions of the same region (Hardoim et al., 2008). Temporal changes in relative frequency of total endophytic fungi were studied (Hardoim et al., 2008). It was established that the plant age, climatic conditions and sampling periods have an impact on plant endophytic colonization. It was stated that matured leaves of teak (*Tectona grandis* L.) and rain tree (*Samanea saman* Merr.) had greater number of genera and species, with higher colonization frequency, than those in the young leaves and their occurrence in leaves increased during rainy season (Hardoim et al., 2008). The endophytic population and frequency tended to differ among sampling dates for all the organs studied, namely, young leaves, petiole, and twigs of *Ginkgo biloba* L (Hardoim et al., 2008). They proved that the occurrence of *Phyllosticta* sp. in both leaves and petioles was first detected in August and peaked in October with none in the month of May. *Phomopsis* sp. was detected in twigs throughout the growing season. These results suggest that the distribution of the two dominant endophytic fungi was organ-specific and differed within seasons. The same results have been observed with *T. esculentum*.

Alphaproteobacteria (Unclassified), *Epsilonbacteria* (*Campylobacter*) and *Fusobacteria* (*Fusobacterium*) were found in leaves as well as in the tubers, but were

absent in stems. However these results were not found in all leaf and root samples. Above ground, *Clostridiales* (*Parvimonas*) were exclusively stem and leave dwellers.

Our result showed a high speciation of the tuber and leaves microbiome compelled by the plant response to specific exterior conditions. It can be suggested that, according to Berg et al. (2014), plant species, cultivar, age, health, and developmental stage are not the only factors that impact on plant microbial communities. There is a myriad of abiotic factors such as the rhizosphere microbiome, soil properties, nutrient status, and climatic conditions that control structural and functional plant endophytic diversity. It has also been suggested that external factors to the host plant such as soil, geographic factors, and human interference contribute to the overall plant endophytic microbial structure and function (Rout & Southworth, 2013).

Moreover, Mano et al. (2008) and Hardoin et al. (2012) detected that the endophytic bacterial community in the leaves of rice plants cultivated in the field was similar to that found in seed tissues. It was drastically different from that inside root tissue. Their results proposed that rice seed endophytes are generally adapted to plant tissue and rapidly colonize rice shoots, in which there is less competition than in the respective root, which is bathed in rich bacterial communities. This would explain the active microbial colonization of *T. esculentum*

5.4.4 *T. esculentum* below ground (tuberous root) endophytic microbiota

Firmicutes were predominantly found in the tuber (below ground) especially dominated by the genus *Bacillus*. *Clostridiales* were exclusively stems and leave dweller. Though belonging to the same phyla, they were clearly separated along plant organ lines (Figure 17 and 18, Table 16). *Proteobacteria*, the second largest phylum was highly dominated by the *Gammaproteobacteria* class of which the *Enterobacteriaceae* family dominated. Eighteen sequences were assigned to the genus *Enterobacter*, 11 to *Pantoea* and one to the genus *Leclercia* and *Enterobacteriaceae* not classified at the genus level. Despite being the inhabitant of human and animal intestines, it was reported that *Enterobacteriaceae* are also indigenous members of several plant microbiomes (Erlacher, Cardinale, Grube, & Berg, 2015). Torres et al. (2008) and Trotel-Aziz, Couderchet, Biagianti, & Aziz Aziz (2008) reported endophytic *Enterobacteriaceae* members in grapevine, sugarcane, cocoa, citrus, *Eucalyptus* and Soybean crop species with an ubiquitous distribution. This suggests that the plant and a given niche could cooperatively shape the structure of their endophytic communities (Berg & Smalla, 2009). A given endophytic microbiome can be modified by factors such as the soil physicochemical environmental conditions, plant growth phase and plant physiological state, as well as by diverse environmental factors (Mercado-Blanco & Lugtenberg, 2014). The fluctuation in physicochemical properties between below and above ground and nutrient supply between these niches would explain the differences in bacterial community structure and the bacterial composition. This would result in the so-called competent endophytes that would thrive in the plant even under

adverse conditions (Hardoim, Hardoim, Overbeek, & Elsas, 2012). This would explain the dominance of *Bacillus* and *Enterobacter* in root. ‘Competent endophyte’ microorganism successfully colonizes the plant tissues and has the capacity to incite plant physiology and be selectively favoured, leading to beneficial maintenance of the plant-microbe association (Hardoim et al., 2012).

Of the bacteria assigned to the *Enterobacteriaceae* family, *Pantoea* and *Enterobacter* species have been described to play an important role in *Zea mays* and wheat (Asis & Adachi, 2004). They have been isolated from potato stems (Asis & Adachi, 2004), rice seeds (Ruppel, Hecht-Buchholz, Remus, Ortmann, & Schmelzer, 1992), and citrus leaves (Araújo et al., 2001) as well.

There are 22 known species in the genera *Pantoea* that have been isolated from different healthy plant parts organs such as aerial surfaces, within tissues (seeds, fruits, roots, tubers and leaves) and even the rhizosphere. This non spore forming gram negative bacterium has also been found in human samples (urine, blood, wounds, and internal organs) and animals (Chauhana, Bagyaraja, Selvakumarb, & Sundaramc, 2015). They are seen as pathogens. However, a pathogenic role in humans has been observed in some instances (Chauhana et al., 2015).

Several researchers have demonstrated that *Pantoea* spp. endophytically is endowed with systemic resistance induction against plant-pathogen microorganisms (Bardina, Huang, Liua, & Yanke, 2003). Successful tests were conducted on canola, sunflower, dry pea, and sugar beet (Bardina et al., 2003). It has been reported that, though

the mode of action has not been elucidated, *Pantoea* has successfully been used as an effective biological control agent both in pre and post harvest stages against disease caused by *Penicillium expansum*, *Botrytis cinerea* and *Rhizopus stolonifer* in apples (*Malus domestica*), against *Penicillium digitatum* and against *Penicillium italicum* in pear (*Pyrus communis*) and Citrus species (Poppe, Vanhoutte, & Hofte, 2003; Nunes et al., 2008).

Similarly, these bacteria can act as plant growth promoters by contributing to the plant nitrogen intake in non-symbiotic associations (Asis & Adachi, 2004), solubilizing phosphorus (Marefat, Ophel-Keller, & McKay, 2004), and stimulating the production of phytohormones (Tsavkelova et al., 2007). They reported that *Pantoea* strains growing under extensive array of abiotic factors such as temperatures, pH and salt concentration. It is responsible for multiple plant growth traits like Indole-3-Acetic Acid (IAA) production, siderophore production, Hydrogen Cyanide (HCN) production (Chauhana et al., 2015).

The poplar tree is known to thrive in marginal soils. When associated with the plant endophyte *Enterobacter sp 638*, this bacterium can improve the growth of poplar trees on marginal soils by as much as 40% (Taghavi et al., 2010). Studies were conducted by Taghavi et al. (2010) Taghavi & Lelie (2013) to identify the plant endophyte *Enterobacter sp 638* specific genes responsible for its symbiotic niche adaptation in Poplar trees. In their work, the existence of genes that code for putative proteins that may be involved in the successful *Enterobacter sp 638* survival in the rhizosphere and responsible for this species to cope with oxidative stress or uptake of nutrients released by plant roots

were suggested. These proteins are believed to be also responsible for promoting the bacteria root adhesion (pili, adhesion, hemagglutinin, and cellulose biosynthesis) and favour its colonization and establishment inside the host plant using chemotaxis, flagella, and cellobiose phosphorylase. The same proteins have been advocated to contribute to plant protection against pathogenic fungal and bacterial infections (siderophore production and synthesis of the antimicrobial compounds 4-hydroxybenzoate and 2-phenylethanol), and improved plant growth and development through the production of the phytohormones indole acetic acid, acetoin, and 2, 3-butanediol. Though these pathways have never been proven to exist in endophytes associated with *T. esculentum*, they are most likely to be taking place given its adaptation to the arid environment.

According to Downie (2010), bacterial protein secretions are critical keys in the plant- bacterium interaction outcome. Consequently, the molecular cross-talk between the partners requires a finely-tuned metabolic coordination and transcriptional regulation. In the same study, metabolite analysis showed that, the release of Volatile Organic Compounds (VOC) acetoin and 2, 3-butanediol in *Arabidopsis* triggered the greatest level of growth promotion and induced resistance. In *Arabidopsis*, interestingly, according to the same author, both the genetic determinants required for sucrose metabolism and the synthesis of acetoin and 2, 3-butanediol are clustered on a genomic island (Taghavi, et al., 2010; Taghavi & van der Lelie, 2013). These results propose a close interaction between *Enterobacter sp.* 638 and its poplar host, where the availability of sucrose, a major plant sugar, affects the synthesis of plant growth promoting phytohormones by the endophytic bacterium.

In all these discourses on plant endophytes and their role and probable origins, focus is made only on their rhizosphereic origin. Little attention has been given to vertically transmitted microbes from the mother to the offspring (Ramírez-Puebla et al., 2013) which might eventually colonize any part of the plant. While studying seed-borne and root endophytic community bacteria associated with the common bean (*Paseolus vulgaris*) under gnotobiotic conditions, López-López, Rogel, Ormeno-Orrillo, Martínez-Romero, & Martínez-Romero (2010) reported the same phyla as the one found in our study with a higher diversity. Their findings supported the notion that seed isolates were *bona fide* endophyte and seed-borne. All 24 genera, at the exception of few (*Arthrobacter*, *Brachybacterium*, and *Enterococcus*), obtained as seed isolates were also recovered from bean roots. Some isolates were present in roots and not in seeds. It was suggested that seed born bacteria were able to multiply and persist in the germinated plant.

This study confirms how aberrant it would be to only consider the rhizosphere as the only source of plant endophytic bacteria. In this study, we can cautiously conclude that at the sight of our community microbial distribution, *T. esculentum* might have developed the ability to select certain microbial consortia for its functional needs. In nutrient limited arid desert ecosystems, plants have developed mechanisms to fully sustain their primary production for their survival. *Proteobacteria* endophytes could be of uttermost importance since they have been found to be among the few phyla endowed with bacteriochlorophyll-based photosynthetic centers (Makhalanyane et al., 2013). Furthermore, a rare bacteria phylum *Gemmatimonadetes* occurring in the Gobi Desert has shown the characteristic to confer photosynthetic capacity to other phyla by gene lateral transfer (Zeng, Feng,

Medová, Dean, & Koblížek, 2014). These phenomena have a high possibility of occurring in *T. esculentum*, hence the importance of such a study.

Firmicutes are known to be the most abundant phylum in the human body and the soil. Despite environmental constraints and interactions with other microorganisms, some bacteria are able to colonize plant compartments with higher frequency than others (Kumar et al., 2011). This suggests a fine tuned discriminative plant system to comply with its needs of the moment (Kumar et al., 2011). In their study on *Bacillus* colonization in soybean, the population of *Bacillus* fluctuated depending on the age of the plant. The density varied from 0 to 80% of the total microbiota. In this study, it is not clear whether *Firmicutes* density in *T. esculentum* does not fluctuate with regard to the plant needs. In *T. esculentum*, *Firmicutes* were largely dominated by the genera *Bacillus*. Reports on *Bacillus* plant colonization remain largely contradictory between studies (Mandić-Mulec & Prosser, 2011).

Garbeva et al. (2003) demonstrated how *Bacilli* were not a major component of the microbial community colonizing potato despite being a tuberous plant. Likewise, *Bacillus* species are believed to be less rhizosphere competent compared to *Pseudomonas* (Kumar, Prakash, & Johri, 2011). This assumption was later confirmed by Inceoğlu, Salles, Overbeek, & van Elsas (2010) whose study on rhizospheric soils of six potato cultivar showed an almost insignificant presence of *Firmicutes* while they were abundant in the bulk soil.

In contrast, studying the same plant (Marques et al., 2014) came to the conclusion

that not only was *Bacillus* the most abundant in the tuber, but was also correlated to the starch content. The root exudate play a crucial role particularly its chemical composition coupled with environmental factors. Moreover, the plant genotype are most likely to influence the composition and diversity of the root associated bacteria. Besides, rhizospheric competence is a necessary prerequisite for PGPR. It comprises of effective root colonization combined with the ability to survive and proliferate along the growing plant roots in the presence of indigenous microbiota over a period of time.

For *T. esculentum*, the environment might have favoured the spore forming bacteria than the *Pseudomonas* that would find it difficult to survive the long dry season of the Kalahari Desert. For more clarification on *T. esculentum* microbial community, the impact of seed originated bacteria should be assessed. It is still believed that for most of bacterial endophytes, their function or ecology within the host plant has not been established. Still, particular bacterial endophytes could vigorously affect the host physiology (Aeron, Kumar, Pandey, & Maheshwari, 2011). Consequently, they can influence the production of phytohormones and/or the modulation of host ethylene levels (Adesemoye & Kloepper, 2009; Colemann-Dar & Tring, 2014). Many more other plant-growth-promoting functions, such as fixation of N₂, solubilisation of inorganic phosphate, provision of micronutrients, promotion of photosynthetic activity, induction of the plant defence system, production of antibiotics, biotransformation of heavy metals and biodegradation of organic pollutants, might contribute to the enhancement of the host fitness (Berg & Smalla, 2009). The effect of these beneficial functions might be drastically improved when plant endophytes establish synergistic interactions with their plant hosts

(Hardoim, Hardoim, van Overbeek, & van Elsas, 2012).

CHAPTER 6: CONCLUSION

There is no common established approach to study the microbial biodiversity of composite environments. The use of more than one method might provide an enhanced global overview of the microbial composition. To date, this is the most comprehensive and the first study to examine *T. esculentum*'s structure and composition bacterial communities through culture dependent and independent using pyrosequencing approaches.

This study presented evidence that time and sampling sites had no impact on the endophytic numerical microbial recruitment of *T. esculentum*'s tuber, leaves and stems in its native environment. However, plant organs were shown to exert an impact on the microbial density distribution. This led to the conclusion that there is a clear niche partitioning microbial density between below (tuberous roots) and above (leaves and stems) ground parts of *T. esculentum*. The tuberous roots were harbouring the highest microbial density.

Culture-dependent and independent techniques used in this study highlighted that the distribution of microbial endophytic community structures were plant organ specific. The above ground and below ground microenvironments in *T. esculentum* harboured distinct community structures. In addition, phyla that were constantly found in *T. esculentum* specific organs across samples are endowed with well-established critical function in microbial communities. Similarly, most sequences obtained have shown similarities with sequences occurring in water stressed environments with well-

established functional capabilities which may be an important factor in the symbiotic relationship. The results attest that *T. esculentum* microbial colonization is not passive at least at a certain developmental stages.

At the genus level, this study's data set is poorly represented. More than 50% of the sequences were termed unclassified. Additionally, sequencing and BLAST analysis of 16S rRNA and *nifH* genes revealed that *T. esculentum* isolates species formed a tight sub-cluster, strongly supported by bootstrap analysis but distinct from their nearest match from the NCBI database. It can be predicted that *T. esculentum* plants contain a reservoir of undiscovered bacterial diversity. These bacteria may be linked to functional genes or characteristics unique to this plant. However, the rarefaction curves suggest further sampling to confirm the microbial endophytic community structures observed in *T. esculentum*. Thus, *T. esculentum*'s endophytic microbial community might be more complex than it was initially thought.

All isolates associated with *T. esculentum* were all positive for *nifH* gene. This indicate that *T. esculentum* supports one of the most diverse diazotrophic communities' studied so far. The harbouring of nitrogen-fixing bacteria within *T. esculentum*'s leaves, stems and tuberos may be an adaptation to the harsh environment in which this plant occurs, further supporting the hypothesis that plant genotype plays a role in determining which bacteria can colonize the host. All these bacteria have not probably been cultivated or identified yet since the respective *nifH* genes had only nucleotide identity between 85 to 100% to previously reported nucleotide sequences known in RDP database.

Comparison of the 16S rRNA amplicon sequences and the culture dependent techniques revealed that the use of the 16S rRNA gene helped to detect comprehensive levels of community structure in *T. esculentum* while metagenomics showed a higher level of diversity. However, very little is known about *T. esculentum* interactions with the microbial community in natural ecosystems. Hence, describing the interaction of these bacteria with the host plant would be of great interest with regards to our knowledge about the plant–microbe interactions and thereafter their economic potential and environmental benefits in agriculture.

CHAPTER 7: RECOMMENDATIONS

Plant organs have been shown to have an impact on the microbial density distribution in *T. esculentum*. The biological significance of microbial counts in the different plant organs, their fluctuation and their potential agronomic implications at different growth stage, namely the seed germination rates, before, during and after flowering stages, seed protein and tuber starch content, need to be investigated further for *T. esculentum* successful domestication. Similarly, this study demonstrated that the recruitment of root, leaf and tuberous root associated bacterial communities by *T. esculentum* is not congruent with the time of sampling and the plant origin. However, the study does not exclude the plant origin and sampling time importance in maintaining the microbial equilibrium in *T. esculentum*.

Both the plant and the surrounding soil environments exert influences on diazotrophs and other microorganisms residing in the rhizosphere and endophytic compartment (Reinhold-Hurek & Hurek, 1998). Soil chemical factors such as pH, nutrients, and soil organic matter act as ecological filters that limit the microbes that have access to the plant rhizosphere and create local species pools. Plant tissue structures, defense systems, and the interactions between plant and microbes enable plants to recruit and select endophytic microbes from rhizosphere microbes, representing a second layer

of ecological filters (Hardoim et al., 2008; Bulgarelli et al., 2013). These successive sets of ecological filters ultimately determine the microbial assemblages associated with *T. esculentum*. However, this needs to be explored further in *T. esculentum*.

The present study demonstrated that *T. esculentum* microbial community diversity was dependent on plant organ in which it occurs. On the other hand, the role that all these microbial community niches might be playing in *T. esculentum* biology requires further investigations. Such studies should focus on the microbial communities' potential role such as the functions that are relevant for plant health, plant growth promotion and plant increased resiliency in this hostile environment. Similarly, in order to expand the spectrum of understanding of possible plant-endophyte interactions, it would be necessary to study *T. esculentum*'s endophytic communities through different growth stages and for an extended period.

In this study, for estimating the phylogeny and taxonomy in microbial communities, the 16S genes were used. 16S rRNA gene copy variation among strains of the same species may occur and consequently inflate diversity estimates. It would lead to a biased interpretation of the results. Even more, through a metagenomic approach, extracted total community DNA amplified with 16S rRNA is hindered to estimating microbial community diversity (DeAngelis et al., 2009).

For future studies, it can be suggested that other housekeeping genes with single copy marker be sought for their present intertaxa sequence variations such as the RNA polymerase B (*rpoS*) gene, the gyrase B (*gyrB*) gene, *recA* gene family, the heat shock

(*dnaK*) gene and the elongation factor Tu (*tuf*) genes or functional gene markers like (*amoA*) that encodes ammonia monooxygenase, a key enzyme in oxidation of ammonia or (*nifH*) encoding the nitrogenase reductase, a key enzyme in nitrogen fixation (Koskinen, 2013).

For a comprehensive elucidation of bacteria community composition and structures associated with *T. esculentum*, DNA analyses targeted not only active microorganisms, but also non-active microorganisms. There is an urgent need to investigate gene expression directly in *T. esculentum* since it can provide a more detailed understanding of the dynamics of the functional population, the activities of specific groups. Transcriptomic /megatranscriptomic measurement would help to shed light on the functional active structure and complexity of *T. esculentum* microbiomes. Further studies should be focused on exploring and testifying the potential roles, such as the functions that are relevant to global N cycling, plant health, and plant growth promotion, of particular phylotypes (e.g., Gammaproteobacteria affiliated families).

Bacterial isolates were identified based on the 16S r RNA sequence homology analysis with existing sequences in the databases (RDP). At the genus level, this study's data set was poorly represented. More than 50% of the sequences were termed unclassified, 49% of accounted taxa showed an identity similarity values below 98.7-99%. This showed how the use of 16S rRNA in this specific case could not clearly provide the taxonomic power to resolve the identity of these species. They might be a high proportion of putatively new species associated with *T. esculentum*. I would suggest manipulation

that would involve DNA-DNA hybridization. This technique is used for proposed new species and for the definitive assignment of a strain with ambiguous properties to the correct taxonomic unit.

REFERENCES

- Acosta-Martinez, V., Dowd, S., Sun, Y., & Allen, V. (2008). Tag-encoded pyrosequencing analysis of bacterial diversity in a single soil type as affected by management and land use. *Soil Biology and Biochemistry*, 40, 2762–2770.
- Adesemoye, A. O., & Kloepper, J. W. (2009). Plant–microbes interactions in enhanced fertilizer-use efficiency. *Applied Microbiology and Biotechnology*, 85(1), 1–12. <http://doi.org/10.1007/s00253-009-2196-0>
- Aeron, A., Kumar, S., Pandey, P., & Maheshwari, D. (2011). Emerging Role of Plant Growth Promoting Rhizobacteria in Agrobiology. In D. Maheshwari (Ed.), *Bacteria in Agrobiology: Crop Ecosystems* (pp. 1–36). Heidelberg/ Dordrecht/ London/ New York: Springer-Verlag Berlin Heidelberg.
- Ahrens, H. (1982). CHATFIELD, C., A. J. COLLINS: Introduction to Multivariate Analysis. *Biometrical Journal*, 24(5), 436.
- Aires, T., Serrão, E. A., Kendrick, G., Duarte, C. M., & Arnaud-Haond, S. (2013). Invasion is a community affair: Clandestine followers in the bacterial community associated to green algae, *Caulerpa racemosa*, track the invasion source. *PLoS ONE*, 8(7), e68429. <http://doi.org/10.1371/journal.pone.0068429>
- Altschul, S., Madden, T., Schaffer, A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acid Research*, 25(17), 3389–3402.
- Amarteifio, J. O., & Moholo, D. (1998). The Chemical Composition of Four Legumes Consumed in Botswana. *Journal of Food Composition and Analysis*, 11, 329–332. <http://doi.org/10.1006/jfca.1998.0595>
- Amonsou, E. O., Taylor, J. R. N., Beukes, M., & Minnaar, A. (2012). Composition of marama bean protein. *Food Chemistry*, 130(3), 638–643. <http://doi.org/10.1016/j.foodchem.2011.07.097>
- Amonsou, E. O., Taylor, J. R. N., & Minnaar, A. (2013). Adhesive potential of marama bean protein. *International Journal of Adhesion and Adhesives*, 41, 171–176. <http://doi.org/10.1016/j.ijadhadh.2012.11.002>
- Andersona, M., & Habigerb, J. (2013). Characterization and Identification of

- Productivity-Associated Rhizobacteria in Wheat. *Applied Environmental Microbiology*, 79(23), 7428–7438.
- Andreote, F., Azevedo, J., & Araújo, W. (2009). Assessing the diversity of bacterial communities associated with plants. *Brazilian Journal of Microbiology*, 40(3), 417–432. <http://doi.org/10.1590/S1517-83822009000300001>
- Andrew, D. R., Fitak, R. R., Munguia-Vega, A., Racolta, A., Martinson, V. G., & Dontsova, K. (2012). Abiotic Factors Shape Microbial Diversity in Sonoran Desert Soils. *Applied and Environmental Microbiology*, 78(21), 7527–7537. <http://doi.org/10.1128/AEM.01459-12>
- Angus, A. a, & Hirsch, A. M. (2015). Biofilm Formation in the Rhizosphere: Multispecies Interactions and Implications for Plant Growth, (MARCH 2013). <http://doi.org/10.1002/9781118297674.ch66>
- Ansorge, W. (2009). Next-generation DNA sequencing techniques. *New Biotechnology*, 25(4), 195–103.
- Araújo, W., Maccheroni, W. J., Aguilar-Vildoso, C., Barroso, P., Saridakis, H., & Azevedo, J. (2001). Variability and interactions between endophytic bacteria and fungi isolated from leaf tissues of citrus rootstocks. *Canadian Journal of Microbiology*, 47(3), 229–236.
- Aravind, R., Kumar, A., Eapen, S. J., & Ramana, K. V. (2009). Endophytic bacterial flora in root and stem tissues of black pepper (*Piper nigrum* L.) genotype: isolation, identification and evaluation against *Phytophthora capsici*. *Letters in Applied Microbiology*, 48(1), 58–64. <http://doi.org/10.1111/j.1472-765X.2008.02486.x>
- Armougom, F., & Raoult, D. (2009). Exploring microbial diversity using 16S rRNA high-throughput methods. *Journal of Computer Sciences and Systematic Biology*, 2, 74–92.
- Artursson, V., Finlay, R., & Jansson, J. (2006). Interactions between arbuscular mycorrhizal fungi and bacteria and their potential for stimulating plant growth. *Environmental Microbiology*, 8(1), 1–10.
- Asis, C. a., & Adachi, K. (2004). Isolation of endophytic diazotroph *Pantoea* agglomerans and nondiazotroph *Enterobacter asburiae* from sweetpotato stem in Japan. *Letters in Applied Microbiology*, 38(1), 19–23. <http://doi.org/10.1046/j.1472-765X.2003.01434.x>

- Atlas, R. ., & Bartha, R. (1998). *Microbial Ecology: Fundamentals & Applications*. (B. Pubs, Ed.) (4th editio). Company Inc. Calif.
- Bagwell, C. E., Dantzler, M., Bergholz, P. W., & Lovell, C. R. (2000). Host-specific ecotype diversity of rhizoplane diazotrophs of the perennial glasswort *Salicornia virginica* and selected salt marsh grasses. *Aquatic Microbial Ecology*, 23, 293–300.
- Bai, Y., Zhou, X., & Smith, D. (2003). Enhanced soybean plant growth resulting from coinoculation of *Bacillus* strains with *Bradyrhizobium japonicum*. *Crop Science*, 43, 177–1781.
- Baldani, J. I., & Baldani, V. L. D. (2005). History on the biological nitrogen fixation research in graminaceous plants: Special emphasis on the Brazilian experience. *Anais Da Academia Brasileira de Ciencias*, 77(3), 549–579. <http://doi.org/10.1590/S0001-37652005000300014>
- Bardina, S. D., Huang, H. C., Liua, L., & Yanke, L. J. (2003). Control, by microbial seed treatment, of dampingoff caused by *Pythium* sp. on canola, safflower, dry pea, and sugar beet. *Canadian Journal of Plant Pathology*, 25(3), 268–275.
- Barracough, T., Hughes, M., Ashford-Hodges, N., & Fujisawa, T. (2009). Inferring evolutionarily significant units of bacterial diversity from broad environmental surveys of single-locus data. *Biology Letters*, 5, 425–428.
- Barrios, E. (2007). Soil biota, ecosystem services and land productivity. *Ecological Economics*, 64(2), 269–285. <http://doi.org/10.1016/j.ecolecon.2007.03.004>
- Barriuso, J., Valverde, J. R., & Mellado, R. P. (2011). Estimation of bacterial diversity using next generation sequencing of 16S rDNA: a comparison of different workflows. *BMC Bioinformatics*, 12(1), 473. <http://doi.org/10.1186/1471-2105-12-473>
- Basualdo, C. V. (2011). Choosing the best non-parametric richness estimator for benthic macroinvertebrates databases. *Revista de La Sociedad Entomológica Argentina*, 70(1-2), 27–38.
- Benedito, V. A., Torres-Jerez, I., Murray, J. D., Andriankaja, A., Allen, S., Kakar, K., ... Udvardi, M. K. (2008). A gene expression atlas of the model legume *Medicago truncatula*. *The Plant Journal*, 55(3), 504–513. <http://doi.org/10.1111/j.1365-313X.2008.03519.x>
- Bent, S., & Forney, L. (2008). The tragedy of the uncommon: understanding limitations

- in the analysis of microbial diversity. *ISME Journal*, 2(7), 689–695.
- Berendsen, R. L., Pieterse, C. M. J., & Bakker, P. A. H. M. (2012). The rhizosphere microbiome and plant health. *Trends in Plant Science*, 17(8), 478–486. <http://doi.org/10.1016/j.tplants.2012.04.001>
- Berg, G., Grube, M., Schlöter, M., & Smalla, K. (2014). Unraveling the plant microbiome: looking back and future perspectives. *Frontiers in Microbiology*, 5, 1–7. <http://doi.org/10.3389/fmicb.2014.00148>
- Berg, G., Mahnert, A., & Moissl-Eichinger, C. (2014). Beneficial effects of plant-associated microbes on indoor microbiomes and human health? *Frontier in Microbiology*, V 5, 1–5.
- Berg, G., Roskot, N., Steidle, A., Eberl, L., Zock, A., & Smalla, K. (2002). Plant-Dependent Genotypic and Phenotypic Diversity of Antagonistic Rhizobacteria Isolated from Different Verticillium Host Plants Plant-Dependent Genotypic and Phenotypic Diversity of Antagonistic Rhizobacteria Isolated from Different Verticillium Host Pl. *Applied and Environmental Microbiology*, 68(7), 3328–3338. <http://doi.org/10.1128/AEM.68.7.3328>
- Berg, G., & Smalla, K. (2009). Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. *FEMS Microbiology Ecology*, 68(1), 1–13.
- Berg, G.; Krechel, A.; Ditz, M.; Sikora, R.A.; Ulrich, J. H. (2005). Endophytic and ectophytic potato associated bacterial communities differ in structure and antagonistic function against plant pathogenic fungi. *FEMS Microbiology Ecology*, 51, 215–229.
- Berlec, A. (2012). Novel techniques and findings in the study of plant microbiota: search for plant probiotics. *Plant Science : An International Journal of Experimental Plant Biology*, 193-194, 96–102. <http://doi.org/10.1016/j.plantsci.2012.05.010>
- Berry, D., & Widder, S. (2014). Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Frontier in Microbiology*, 5(219), 1–14.
- Bisseling, T., Dangl, J., & Schulze-Lefert, P. (2009). Next-Generation Communication. *Science*, 5928, 691.
- Bodenhause, N., Horton, M. W., & Bergelson, J. (2013). Bacterial Communities

- Associated with the Leaves and the Roots of *Arabidopsis thaliana*. *PLoS ONE*, 8(2), e56329. <http://doi.org/10.1371/journal.pone.0056329>
- Bohannon, B. J., & Hughesy, J. (2003). New approaches to analyzing microbial biodiversity data. *Current Opinion in Microbiology*, 6, 282–287.
- Bordenstein, S., & Reznikoff, W. (2005). Mobile DNA in obligate intracellular bacteria. *Nature Review Microbiology*, 3(9), 688–99.
- Borer, E.T., Kinkel, L. L., May, G., & Seabloom, E. W. (2013). “The world within: quantifying the determinants and outcomes of a host’s microbiome,.” *Basic and Applied Ecology*, 14(2013), 533.
- Bower, N., Hertel, K., Oh, J., & Storey, R. (1988). Nutritional evaluation of Marama bean (*Tylosema esculentum*, Fabaceae): analysis of the seed. *Economic Botany*, 42(4), 533–540.
- Bright, M., & Bulgheresi, S. (2010). A complex journey: transmission of microbial symbionts. *Nature Review Microbiology*, 8(3), 218–230.
- Bueche, M., Wunderlin, T., Roussel-Delif, L., Thomas Junier, L., Junier, S., Jeanneret, N., & Pilar. (2013). Quantification of Endospore-Forming Firmicutes by Quantitative PCR with the Functional Gene *spo0A*. *Applied and Environmental Microbiology*, 79(17), 5302–5312.
- Bulgarelli, D., Schlaeppi, K., Spaepen, S., van Themaat, E. V. L., & Schulze-Lefert, P. (2013). Structure and Functions of the Bacterial Microbiota of Plants. *Annual Review of Plant Biology*, 64(1), 807–838. <http://doi.org/10.1146/annurev-arplant-050312-120106>
- Burbano, C., Grönmeyer, J., Hurek, T., & Reinhold-Hurek, B. (2015). Microbial community structure and functional diversity of nitrogen-fixing bacteria associated with *Colophospermum mopane*. *FEMS Microbiology Ecology*, 91(4), fiv030.
- Bürgmann, H., Meier, S., Bunge, M., Widmer, F., & Zeyer, J. (2005). Effects of model root exudates on structure and activity of a soil diazotroph community. *Environmental Microbiology*, 7(1), 1711–1724.
- Burgmann, H., Widmer, F., Sigler, W. ., & Zeyer, J. (2004). New molecular screening tools for the analysis of free-living diazotrophs in soil. *Applied and Environmental Microbiology* 70: 240–247., 70:, 240–247.
- Burns, J., Zehr, J., & Capone, D. (2002). Nitrogen-fixing phylotypes of Chesapeake Bay

- and Neuse River estuary sediments. *Microbial Ecology*, 44(4), 336–343.
- Cai, Y., & Sun, Y. (2011). ESPRIT-Tree: hierarchical clustering analysis of millions of 16S rRNA pyrosequences in quasilinear computational time. *Nucleic Acids Research*, 39(14), e95.
- Cannon, S. B., May, G. D., & Jackson, S. A. (2009). Three Sequenced Legume Genomes and Many Crop Species: Rich Opportunities for Translational Genomics. *Plant Physiology*, 151(3), 970–977. <http://doi.org/10.1104/pp.109.144659>
- Canton, M., Anschutz, P., Coynela, A., Polsenaere, P., Aubry, I., & Poirier, D. (2012). Nutrient export to an Eastern Atlantic coastal zone: first modeling and nitrogen mass balance. *Biogeochemistry*, 107(1-3), 361–377.
- Cardinale, M., Grube, M., Erlicher, A., Quehenberger, J., & Berg, G. (2015). Bacterial networks and co-occurrence relationships in the lettuce root microbiota. *Environmental Microbiology*, 17(1), 239–252.
- Caro-Quintero, A Konstantinidis, K. (2012). Bacterial species may exist, metagenomics reveal. *Environmental Microbiology*, 14(2), 347–355.
- Casas, C., Omacini, M., Montecchia, M. S., & Correa, O. S. (2011). Soil microbial community responses to the fungal endophyte *Neotyphodium* in Italian ryegrass. *Plant Soil*, 340(1), 347–355.
- Castro, S., Silveira, P., Coutinho, A. P., & Figueiredo, E. (2005). Systematic studies in *Tylosema* (Leguminosae). *Botanical Journal of the Linnean Society*, 147(1), 99–115. <http://doi.org/10.1111/j.1095-8339.2005.00353.x>
- Chao, A. (1984). Non parametric estimation of the number of classes in a population? *Scandinavian Journal of Statistics*, 11, 265–270.
- Chauhana, H., Bagyaraja, D. ., Selvakumar, G., & Sundaram, S. . (2015). Novel plant growth promoting rhizobacteria—Prospects and potential. *Applied Soil Ecology*, 95, 38–53.
- Chazdon, R. ., Colwell, R. K., Denslow, J. S., & Guariguata, M. R. (1998). Statistical methods for estimating species richness of woody regeneration in primary and secondary rain forests of Northeastern Costa Rica. *Man and the Biosphere Series*. In J. A. Dallmeier, F., Comiskey (Ed.), *Forest biodiversity research, monitoring and modeling: conceptual background and old world case studies* (pp. 285–309).
- Chen, W., Zhang, C., Cheng, Y., Zhang, S., & Zhao, H. (2013). A comparison of

- methods for clustering 16S rRNA sequences into OTUs. *PloS One*, 8(8), e70837.
- Chen, W.-M., Faria, S. M. de, Stralioetto, R., Pitard, R. M., Simões-Araújo, J. L., Chou, J.-H., ... Euan, K. (2005). Proof that Burkholderia Strains Form Effective Symbioses with Legumes: a Study of Novel Mimosa-Nodulating Strains from South America. *APPLIED AND ENVIRONMENTAL MICROBIOLOGY*, 71(11), 7461–7471.
- Chen, X., Eric S, S., Xiang, Y., Kachman, M., Andrews, P. C., & Yanzhuang, W. (2010). Quantitative proteomics analysis of cell cycle regulated Golgi disassembly and reassembly. *The American Society for Biochemistry and Molecular Biology*, in press(in press), in press.
- Chimwamurombe, P., Bruce, T., Foyer, C., Halford, N., Keys, A., Kunert, K., ... Russell, G. (2009). Domestication of *Tylosema esculentum* (marama bean) as a crop for Southern Africa: genetic diversity of the Omitara marama subpopulation of Namibia. *Aspects of Applied Biology*, 96, 37–43.
- Chingwaru, W., Majinda, R. T., Yeboah, S. O., Jackson, J. C., Kapewangolo, P. T., Kandawa-Schulz, M., & Cencic, A. (2011). *Tylosema esculentum* (Marama) Tuber and Bean Extracts Are Strong Antiviral Agents against Rotavirus Infection. *Evidence-Based Complementary and Alternative Medicine : eCAM*, 2011, 284795. <http://doi.org/10.1155/2011/284795>
- Chowdhury, S. P., Schmid, M., Hartmann, A., & Tripathi, A. K. (2009). Diversity of 16S-rRNA and nifH genes derived from rhizosphere soil and roots of an endemic drought tolerant grass, *Lasiurus sindicus*. *European Journal of Soil Biology*, 45(1), 114–122. <http://doi.org/10.1016/j.ejsobi.2008.06.005>
- Coelho, M., de Vos, M., Carneiro, N., Marriel, I., Paiva, E., & Seldin, L. (2008). Diversity of nifH gene pools in the rhizosphere of two cultivars of sorghum (*Sorghum bicolor*) treated with contrasting levels of nitrogen fertilizer. *FEMS Microbiology Letters*, 279, 15–22.
- Cole, J. R., Qiong, W., Jordan A, F., Benli, C., , Donna M, M., , Yanni, S., ... James M., T. (2014). Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Research*, 42, D633–D642.
- Cole, J., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R., ... Tiedje, J. (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Research*, 37, D141–D145.

- Colemann-Dar, D., & Tring, S. (2014). Building the crop of tomorrow: advantage of symbiont-based approach to improving abiotic stress tolerance. *Frontier in Microbiology*, 5, 1–6.
- Compant, S., Clément, C., & Sessitsch, A. (2010). Plant growth-promoting bacteria in the rhizo- and endosphere of plants: Their role, colonization, mechanisms involved and prospects for utilization. *Soil Biology and Biochemistry*, 42(5), 669–678. <http://doi.org/10.1016/j.soilbio.2009.11.024>
- Compant, S., Duffy, B., Nowak, J., Cle, C., & Barka, E. a. (2005). Use of Plant Growth-Promoting Bacteria for Biocontrol of Plant Diseases : Principles , Mechanisms of Action , and Future Prospects MINIREVIEW Use of Plant Growth-Promoting Bacteria for Biocontrol of Plant Diseases : Principles , Mechanisms of Action , and. *Applied and Environmental Microbiology*, 71(9), 4951–4959. <http://doi.org/10.1128/AEM.71.9.4951>
- Compant, S., Van Der Heijden, M. G. A., & Sessitsch, A. (2010). Climate change effects on beneficial plant–microorganism interactions. *FEMS Microbiology Ecology*, 73(2), 197–214.
- Conn, V., Walker, A., & Franco, C. (2008a). Endophytic Actinobacteria induce defence pathways in *Arabidopsis thaliana*. *Molecular Plant-Microbe Interactions*, 21(2), 208–218.
- Conn, Walker, R., & Franco, C. M. M. (2008b). in *Arabidopsis thaliana*, 21(2), 208–218.
- Craig, S., Kannadan, S., Flory, S. L., Seifert, E. K., Whitney, K. D., & Rudgers, J. A. (2011). Potential for endophyte symbiosis to increase resistance of the native grass *Poa alsodes* to invasion by the non-native grass *Microstegium vimineum*. *Symbiosis*, 53(17), 17–28.
- Curtis, T., Head, I., Lunn, M., Woddckock, S., Schloss, P., & Sloan, W. (2006). What is the extent of prokaryotic diversity? *Philosophical Transaction of the Royal Society B*, 361, 2023–2037.
- Dai, J., Liu, X., & Wang, Y. (2013). Diversity of endophytic bacteria in *Caragana microphylla* grown in the Desert grassland of Ningxia Hui Autonomous Region of China. *Genetic and Molecular Research*, 13(2), 2349–2358.
- Dakora, F., Beane, A., Thomas, T., Asiwe, J., Naab, J., P, T. ., & Chimphango, S. (2009). Identifying useful biological trait in indigenous African legumes. In Z. Tadele (Ed.), *New approaches to plant breeding of orphans crops in Africa* (p.

1921). Bern.

- Dakora, F. D., & Keya, S. O. (1997). Contribution of legume nitrogen fixation to sustainable agriculture in Sub-Saharan Africa. *Soil Biology and Biochemistry*, 29(5-6), 809–817. [http://doi.org/10.1016/S0038-0717\(96\)00225-8](http://doi.org/10.1016/S0038-0717(96)00225-8)
- Dakora, F., Lawlor, D., & Sibuga, K. (1999). Assessment of symbiotic nitrogen nutrition in marama bean (*Tylosema esculentum* L.), a tuber-producing underutilized African grain legume. *Symbiosis*, 27(3-4), 269 – 277.
- Dale, C; Moran, N. (2006). Molecular interactions between bacterial symbionts and their hosts. *Cell*, 126(3), 453–465.
- DeAngelis, K. M., Brodie, E. L., DeSantis, T. Z., Andersen, G. L., Lindow, S. E., & Firestone, M. K. (2009). Selective progressive response of soil microbial community to wild oat roots. *Isme Journal*, 3(2), 168–178. <http://doi.org/DOI 10.1038/ismej.2008.103>
- Dedysh, S. N., Ricke, P., & Liesack, W. (2004). NifH and NifD phylogenies: an evolutionary basis for understanding nitrogen fixation capabilities of methanotrophic bacteria. *Microbiology*, 150, 1301–1313.
- Delseny, M., Han, B., & Hsing, Y. (2010). High throughput DNA sequencing: The new sequencing revolution. *Plant Science*, 179(4), 407–422.
- Deslippe, J. R., & Egger, K. N. (2006). Molecular diversity of nifH genes from bacteria associated with high arctic dwarf shrubs. *Microbial Ecology*, 51(4), 516–25. <http://doi.org/10.1007/s00248-006-9070-8>
- Di Bella, J. M., Bao, Y., Gloor, G. B., Burton, J. P., & Reid, G. (2013). High throughput sequencing methods and analysis for microbiome research. *Journal of Microbiological Methods*, 95(3), 401–414. <http://doi.org/10.1016/j.mimet.2013.08.011>
- Di Bonaventura, M., Lee, E., DeSalle, R., & Planet, P. (2010). A whole-genome phylogeny of the family Pasteurellaceae. *Molecular Phylogenetics and Evolution*, 54, 950–956.
- Ding, Y., Wang, J., Liu, Y., & Chen, S. (2005). Isolation and identification of nitrogen-fixing bacilli from plant rhizospheres in Beijing region. *Journal of Applied Microbiology*, 99(5), 1271–1281.
- Dixon, R., & Kahn, D. (2004). Genetic regulation of biological nitrogen fixation. *Nature*

- Reviews Microbiology*, 2, 621–631.
- Doty, S. L., Oakley, B., Xin, G., Kang, J. W., Singleton, G., Khan, Z., ... Staley, J. T. (2009). Diazotrophic endophytes of native black cottonwood and willow. *Symbiosis*, 47(1), 23–33. <http://doi.org/10.1007/BF03179967>
- Douglas, A. (1998). Host benefit and the evolution of speciation in symbiosis. *Heredity*, 81, 599–603.
- Downie, S. R., Spalik, K., Katz-Downie, D. S., & Reduron, J.-P. (2010). Major clades within Apiaceae subfamily Apioideae as inferred by phylogenetic analysis of nrDNA ITS sequences. *Plant Diversity and Evolution*, 128(1), 111–136. <http://doi.org/10.1127/1869-6155/2010/0128-0005>
- Droge, B., Doré, H., Borland, S., Wisniewski-Dyé, F., & Prigent-Combaret, C. (2012). Which specificity in cooperation between phytostimulating rhizobacteria and plants? *Research in Microbiology*, 163(8), 500–510. <http://doi.org/10.1016/j.resmic.2012.08.006>
- Drown, D., Zee, P. C., Brandvain, Y., & Wade, M. J. (2013). Evolution of transmission mode in obligate symbionts. *Evolution Ecological Research*, 15, 43–59.
- Dudeja, S. S., Giri, R., Saini, R., Suneja-Madan, P., & Kothe, E. (2012). Interaction of endophytic microbes with legumes. *Journal of Basic Microbiology*, 52(3), 248–260. <http://doi.org/10.1002/jobm.201100063>
- Dumont, M., Lüke, C., Deng, Y., & Frenzel, P. (2014). Classification of pmoA amplicon pyrosequences using BLAST and the lowest common ancestor method in MEGAN. *Frontier in Microbiology*, 5(34), 1–11.
- Duranti, M; Gius, C. (1997). Legume seeds: protein content and nutritional value. *Field Crop Research*, 53(1), 31–45.
- Edwards, J., Johnson, C., Santos-Medellín, C., Lurie, E., Podishetty, N. K., Bhatnagar, S., ... Jeffery, L. D. (2015). Structure, variation, and assembly of the root-associated microbiomes of rice. *Proceedings of the National Academy of Sciences of the United States of America*, 112(8), E911–E920. <http://doi.org/10.1073/pnas.1414592112>
- El-Deeb, B., Fayez, K., & Gherbawy, Y. (2012). medicinal plant in Saudi Arabia desert and their antimicrobial activities. *Journal of Plant Interactions*, 9145(June 2015), 1–9. <http://doi.org/10.1080/17429145.2012.680077>

- Emiliani, G., Mengoni, A., Maida, I., Perrin, E., Chiellini, C., Fondi, M., ... Fani, R. (2014). Linking Bacterial Endophytic Communities to Essential Oils: Clues from *Lavandula angustifolia* Mill. *Evidence-Based Complementary and Alternative Medicine : eCAM*, 2014, 650905. <http://doi.org/10.1155/2014/650905>
- Eppelleberg, I., & Fedor, P. (2003). Attribute to Claude Shannon (1916-2001) and plea for more rigorous use of species richness, species diversity and the Shannon-Wiener's Index. *Global Ecology & Biogeography*, 12, 177–179.
- Erlacher, A., Cardinale, M., Grube, M., & Berg, G. (2015). Biotic Stress Shifted Structure and Abundance of Enterobacteriaceae in the Lettuce Microbiome. *PLOS One*, 10(2), e0118068.
- Ernst, M., Mendgen, K. W., & Wirsal, S. G. R. (2003). Endophytic fungal mutualists: seed-borne *Stagonospora* spp. enhance reed biomass production in axenic microcosms. *Molecular Plant-Microbe Interactions : MPMI*, 16(7), 580–587. <http://doi.org/10.1094/MPMI.2003.16.7.580>
- Eschenhagen, M., Schuppler, M., & Röske, I. (2003). Molecular characterization of the microbial community structure in two activated sludge systems for the advanced treatment of domestic effluents. *Water Research*, 37(13), 3224–3232.
- Fakruddin, M., & Mannan, K. S. Bin. (2013). Methods for Analyzing Diversity of Microbial Communities in Natural Environments. *Ceylon Journal of Science (Biological Sciences)*, 42(1), 19–33. <http://doi.org/10.4038/cjsbs.v42i1.5896>
- Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, 39, 783–791.
- Ferrando, L., Fernández Mañay, J., & Fernández Scavino, A. (2012). Molecular and culture-dependent analyses revealed similarities in the endophytic bacterial community composition of leaves from three rice (*Oryza sativa*) varieties. *FEMS Microbiology Ecology*, 80(3), 696–708. <http://doi.org/10.1111/j.1574-6941.2012.01339.x>
- Ferreira, A., Quecine, M. C., Lacava, P. T., Oda, S., Azevedo, J. Lã., & Araújo, W. L. (2008). Diversity of endophytic bacteria from *Eucalyptus* species seeds and colonization of seedlings by *Pantoea agglomerans*. *FEMS Microbiology Letters*, 287(1), 8–14. <http://doi.org/10.1111/j.1574-6968.2008.01258.x>
- Fierer, N., Leff, J. W., Adams, B. J., Nielsen, U. N., Bates, S. T., Lauber, C. L., ... Caporaso, J. G. (2012). Cross-biome metagenomic analyses of soil microbial

- communities and their functional attributes. *Proceedings of the National Academy of Sciences*, 109(52), 21390–21395. <http://doi.org/10.1073/pnas.1215210110>
- Fierer, N., & Lennon, J. (2011). The generation and maintenance of diversity in microbial communities. *American Journal of Botany*, 98(3), 439–48.
- Fischer, S. E., Fischer, S. I., Magris, S., & Mori, G. B. (2007). Isolation and characterization of bacteria from the rhizosphere of wheat. *World Journal of Microbiology and Biotechnology*, 23(7), 895–903. <http://doi.org/10.1007/s11274-006-9312-4>
- Foster, J. A., Bunge, J., Gilbert, J. A., & Moore, J. H. (2012). Measuring the microbiome: perspectives on advances in DNA-based techniques for exploring microbial life. *Brief Bioinform.* 2012 Jul; 13(4):, 13(4), 420–429.
- Fredriksson, N. J., Hermansson, M., & Wilén, B.-M. (2014). Impact of T-RFLP data analysis choices on assessments of microbial community structure and dynamics. *BMC Bioinformatics*, 15, 1–18.
- Fu, H., Wei, Y., Zou, Y., Li, M., Wang, F., Chen, J., ... Ding, L. (2014). Research progress on actinomycetes *Arthrobacter*. *Advance in Microbiology*, 4, 747–753.
- Gaby, J. C., & Buckley, J. C. (2014). A comprehensive aligned nifH gene database: a multipurpose tool for studies of nitrogen-fixing bacteria. *The Journal of Biological Database and Curation*, 2014, bau001.
- Gachon, C., Mingam, A., & Charrier, B. (2004). Real-time PCR: what relevance to plant studies? *Journal of Experimental Botany*, 55(402), 1445–1454.
- Gandure, J; Ketlogetswe, C; Temu, A. (2014). Performance and Emission Characteristics of Tylosema Esculentum Biodiesel in a Diesel Engine: An Experimental Investigation. *Journal of Energy and Power Engineering*, 8, 50–58.
- Garbeva, P., Van Veen, J. a., & Van Elsas, J. D. (2003). Predominant *Bacillus* spp. in agricultural soil under different management regimes detected via PCR-DGGE. *Microbial Ecology*, 45, 302–316. <http://doi.org/10.1007/s00248-002-2034-8>
- Gergen, I., & Harmanescu, M. (2012). Application of principal component analysis in the pollution assessment with heavy metals of vegetable food chain in the old mining areas. *Chemistry Central Journal*, 6(156), 1–13.
- Glick, B. R. (2012). Plant growth-promoting bacteria: mechanisms and applications. *Scientifica*, 2012, 963401. <http://doi.org/10.6064/2012/963401>

- Gomes, N. C. M., & Cleary, D. F. R. (2013). Impact of Mangrove Roots on Bacterial Composition. *Molecular Microbial Ecology of the Rhizosphere*, 1083 – 1088.
- Gougjal, Y., Toumatia, O., Sabaou, N., Barakate, M., Mathieu, F., & Zitouni, A. (2013). Endophytic actinomycetes from spontaneous plants of Algerian Sahara: indole-3-acetic acid production and tomato plants growth promoting activity. *World Journal of Microbiology and Biotechnology*, 29(10), 1821–1829.
- Grönemeyer, J. L., Burbano, C. S., Hurek, T., & Reinhold-Hurek, B. (2011). Isolation and characterization of root-associated bacteria from agricultural crops in the Kavango region of Namibia. *Plant and Soil*, 356(1-2), 67–82.
<http://doi.org/10.1007/s11104-011-0798-7>
- Grünwald, N. J., Goodwin, S. B., Milgroom, M. G., & Fry, W. E. (2003). Analysis of genotypic diversity data for populations of microorganisms. *Phytopathologist*, 93, 738–746.
- Gulumbe, S. U., Dikko, H. G., & Bello, Y. (2013). Analysis of Crime Data using Principal Component Analysis: A case study of Katsina State. *CBN Journal of Applied Statistics*, 3(2), 39–49.
- Gupta, G., Panwar, J., & Jha, P. N. (2013). Natural occurrence of *Pseudomonas aeruginosa*, a dominant cultivable diazotrophic endophytic bacterium colonizing *Pennisetum glaucum* (L.) R. Br. *Applied Soil Ecology*, 64, 252–261.
<http://doi.org/10.1016/j.apsoil.2012.12.016>
- Haiyambo, D. H., Reinhold-Hurek, B., & Chimwamurombe, P. M. (2015). Effects of plant growth promoting bacterial isolates from Kavango on the vegetative growth of *Sorghum bicolor*. *African Journal of Microbiology Research*, 9(10), 725–729.
- Hall, T. A. (1999). BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, 95–98.
- Hammer, Ø., Harper, D. A. T., & Ryan, P. D. (2001). PAST: PALEONTOLOGICAL STATISTICS SOFTWARE PACKAGE FOR EDUCATION AND DATA ANALYSIS. *Palaeontologia Electronica*, 4(1), 1–9.
- Hampp, R., Hartmann, A., & Nehls, U. (2012). Growth and Defence in Plants, 220, 111–139. <http://doi.org/10.1007/978-3-642-30645-7>
- Hardoim, P. R., Andreote, D. F., Reinhold-Hurek, B., Sessitsch, A., Leonard Simon,

- van O., & Jan Dirk, van E. (2011). Rice root-associated bacteria: insights into community structures across 10 cultivars. *FEMS Microbiology Ecology*, 77, 154–164.
- Hardoim, P. R., Hardoim, C. C. P., Overbeek, L. S. van, & Elsas, J. D. van. (2012). Dynamics of Seed-Borne Rice Endophytes on Early Plant Growth Stages. *PLoS ONE*, 7(2), e30438.
- Hardoim, P. R., van Overbeek, L. S., & Elsas, J. D. van. (2008). Properties of bacterial endophytes and their proposed role in plant growth. *Trends in Microbiology*, 16(10), 463–471. <http://doi.org/10.1016/j.tim.2008.07.008>
- Hartley, M. L., Tshamekeng, E., & Thomas, S. M. (2002). Functional heterostyly in *Tylosema esculentum* (Caesalpinioideae). *Annals of Botany*, 89, 67–76. <http://doi.org/10.1093/aob/mcf006>
- Hartmann, A., Rothballer, M., Hense, B. A., & Schröder, P. (2014). Bacterial quorum sensing compounds are important modulators of microbe-plant interactions. *Frontier in Plant Science*, 5, 131.
- Hartmann, A., Schmid, M., Tuinen, D. van, & Berg, G. (2009). Plant-driven selection of microbes. *Plant and Soil*, 321(1-2), 235–257. <http://doi.org/10.1007/s11104-008-9814-y>
- Hill, T., Walsh, K., Harris, J., & Moffet, B. (2003). Using ecological diversity measures with bacterial communities. *FEMS Microbiology Ecology*, 43, 1–11.
- Hirsch, P. R., & Mauchline, T. H. (2012). Who's who in the plant root microbiome? *Nature Biotechnology*, 30, 961–962.
- Hirsch, R., P., Mauchline, H., T., & Clark, I. M. (2010). Culture-independent molecular techniques for soil microbial ecology. *Soil Biology and Biochemistry*, 42(6), 878–887. <http://doi.org/10.1016/j.soilbio.2010.02.019>
- Hoffman, M., & Arnold, A. (2010). Bacteria inhabit living hyphae of phylogenetically diverse fungal endophytes. *Applied Environmental Microbiology*, 76(12), 4063–4075.
- Hollants, J., Leliaert, F., Verbruggen, H., Willems, A., & De Clerck, O. (2013). Permanent residents or temporary lodgers: characterizing intracellular bacterial communities in the siphonous green alga *Bryopsis*. *Proceedings of the Royal Society B*, 280(1754), 1–8.

- Holse, M., Husted, S., & Hansen, Å. (2010). Chemical composition of marama bean (*Tylosema esculentum*)-A wild African bean with unexploited potential. *Journal of Food Composition and Analysis*, 23(6), 648–657.
<http://doi.org/10.1016/j.jfca.2010.03.006>
- Hughes, J., & Bohannan, B. J. (2004). Application of ecological diversity statistics in Microbial Ecology. In J. Hughes & B. J. Bohannan (Eds.), *Molecular Microbial ecology Manual* (pp. 1321–1344). Dordrecht/Boston/London: Kluwer Academic Publisher.
- Hurek, T. B., Reinhold-Hurek, B., Montagu, M. B., & Kellenberger, E. (1994). Root colonization and systematic spreading of *Azoarcus* sp strain BH72 in grasses. *Journal of Bacteriology*, 176, 1913–1923.
- Huse, S., Huber, J., Morrison, H., Sogin, M., & Welch, D. (2007). Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biology*, 8(7), R143.
- Idris, E., Bochow, H., Ross, H., & Boriss, F. (2004). Use of *Bacillus subtilis* as biocontrol agent. 6. Phytohormone action of culture filtrate prepared from plant growth promoting *Bacillus amyloliquefaciens* FZB24, FZB42, FZB45 and *Bacillus subtilis* FZB37. *Journal of Plant Diseases and Protection*, 111, 583–597.
- Inceoğlu, Ö., Al-Soud, W. A., Salles, J. F., Semenov, A. V., & van Elsas, J. D. (2011). Comparative Analysis of Bacterial Communities in a Potato Field as Determined by Pyrosequencing. *PLoS ONE*, 6(8), e23321.
<http://doi.org/10.1371/journal.pone.0023321>
- Inceoğlu, Ö., Salles, J. F., Overbeek, L. van, & van Elsas, J. D. (2010). Effects of Plant Genotype and Growth Stage on the Betaproteobacterial Communities Associated with Different Potato Cultivars in Two Fields. *Applied and Environmental Microbiology* June 2010 Vol. 76 No. 11, 76(11), 3675–3684.
- Iniguez, A., Dong, Y., & Triplett, E. (2004). Nitrogen fixation in wheat provided by *Klebsiella pneumoniae* 342. *Molecular Plant– Microbe Interaction*, 17(10), 1078–1085.
- James, K., & Olivares, F. (1997). Infection and colonization of sugarcane and other Gramineous plants by endophytic diazotrophs. Critical Review in Plant Sciences 17(1): *Critical Review in Plant Sciences*, 17(1), 77–119.
- Jennifer, L., Kirka, A., Beaudette, L., Hartb, M., Moutoglisc, P., John, N. K., ... Trevorsa. (2004). Methods of studying soil microbial diversity. *Journal of*

Microbiological Methods, 58(2), 169–188.

- Jensen, S., K uhl, M., Glud, R., J rgensen, L., & Priem e, A. (2005). Oxic microzones and radial oxygen loss from roots of *Zostera marina*. *Marine Ecology Progress Series* 293: 49–58., 293, 49–58.
- Jiang, Y., Han, L., Chen, X., Yin, M., Zheng, D., Wang, Y., ... Huang, X. (2013). Diversity and Bioactivity of Cultivable Animal Fecal Actinobacteria. *Advances in Microbiology*, 3, 1–13.
- Jorquera, M., Shaharoon, B., Nadeem, S., Mora M, de la L., & Crowley, D. (2012). Plant growth-promoting rhizobacteria associated with ancient clones of creosote bush (*Larrea tridentata*). *Microbial Ecology*, 64(4), 1008–1017.
- Jost, L. (2010). Independence of alpha and beta diversities. *Ecology*, 91(7), 1969–1974.
- Just , A., Thomma, B., & Ievens, L. (2008). Recent advances in molecular techniques to study microbial communities in food-associated matrices and processes. *Food Microbiology*, 25(6), 745–61.
- Kaga, A., Shimizu, T., Watanabe, S., Tsubokura, Y., Katayose, Y., Harada, K., ... Tomooka, N. (2012). Evaluation of soybean germplasm conserved in NIAS genebank and development of mini core collections. *Breeding Science*, 61(5), 566–592. <http://doi.org/10.1270/jsbbs.61.566>
- Kawaguchi, M., & Minamisawa, K. (2010). Plant-Microbe Communications for Symbiosis. *PLANT AND CELL PHYSIOLOGY*, 51(9), 1377–1380.
- Kayitesi, E., De Kock, H. L., Minnaar, A., & Duodu, K. G. (2012). Nutritional quality and antioxidant activity of marama-sorghum composite flours and porridges. *Food Chemistry*, 131, 837–842. <http://doi.org/10.1016/j.foodchem.2011.09.055>
- Keenan, M., Rudgers, M., & Jennifer, A. (2008). Balancing multiple mutualists: asymmetric interactions among plants, arbuscular mycorrhizal fungi, and fungal endophytes. *Oikos*, 117, 310–320.
- Kennedy, I. R., Choudhury, A. T. M. ., & Kecsk s, M. L. (2004). Non-symbiotic bacterial diazotrophs in crop-farming systems: can their potential for plant growth promotion be better exploited? *Soil Biology and Biochemistry*, 36(8), 1229–1244. <http://doi.org/10.1016/j.soilbio.2004.04.006>
- Kernaghan, G. (2013). Functional diversity and resource partitioning in fungi associated with the fine feeder roots of forest trees. *Symbiosis*, 61,(3), 113–123.

- Keshri, J., Mishra, A., & Jha, B. (2013). Microbial population index and community structure in saline–alkaline soil using gene targeted metagenomics. *Microbiological Research*, 168(3), 165–173.
- Kim, D., Geo, P., Cole, T., Pimentel, H., Kelley, R., & Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology* 2013, 14:R36, 14(4), R 36.
- Kim, S. B. (2005). Bacterial Communities Associated with Sand Dune Plants and Their Potential for Plant Growth Promoting Activity. In *International meeting of the federation of korean Microbiological society* (pp. S16–3).
- Kitts, C. L. (2001). Terminal restriction fragment patterns: a tool for comparing microbial communities and assessing community dynamics. *Current Issues in Intestinal Microbiology*, 2(1), 17–25.
- Kizilovaa, A., Titova, L., Kravchenko, I., & Iutinskaya, G. (2012). Evaluation of the Diversity of NitrogenFixing Bacteria in Soybean Rhizosphere by nifH Gene Analysis. *Microbiology*, 81(5), 621–629.
- Knelman, J. E., Legg, T. M., O'Neill, S. P., Washenberger, C. L., González, A., Cleveland, C. C., & Nemergut, D. R. (2012). Bacterial community structure and function change in association with colonizer plants during early primary succession in a glacier forefield. *Soil Biology and Biochemistry*, 46, 172–180. <http://doi.org/10.1016/j.soilbio.2011.12.001>
- Köberl, M., Schmidt, R., Ramadan, E. M., Bauer, R., & Berg, A. G. (2013). The microbiome of medicinal plants: diversity and importance for plant growth, quality and health. *Front Microbiol.* 2013; 4: 400., 4, 1–9.
- Konopka, A. (2009). What is microbial community ecology? *International Society for Microbial Ecology Journal*, 3, 1223–1230.
- Koskinen, K. (2013). *Characterisation of diverse microbial communities and application of novel detection techniques.*
- Koskinen, K., Hultman, J., Paulin, L., Auvinen, P., & Kankaanpää, H. (2011). Spatially differing bacterial communities in water columns of the northern Baltic Sea. *FEMS Microbiology Ecology*, 75, 99–110. <http://doi.org/10.1111/j.1574-6941.2010.00987.x>
- Kozyrovska, N. O. (2013). Crosstalk between endophytes and a plant host within

- information-processing networks. *Biopolymers and Cell*, 29(3), 234–243.
<http://doi.org/10.7124/bc.00081D>
- Kuklinsky-Sobral, J., Araújo, W. L., Mendes, R., Geraldi, I. O., Pizzirani-Kleiner, A. A., & Azevedo, J. L. (2004). Isolation and characterization of soybean-associated bacteria and their potential for plant growth promotion. *Environmental Microbiology*, 6(12), 1244–1251. <http://doi.org/10.1111/j.1462-2920.2004.00658.x>
- Kumar, A., Prakash, A., & Johri, B. N. (2011). *Bacillus* as PGPR in Crop Ecosystem. In D. K. Maheshwari (Ed.), *Bacteria in Agrobiolgy: Crop Ecosystems*, (p. 434). Verlag Berlin Heidelberg: Springer-Verlag Berlin Heidelberg.
- Laguerre, G., Nour, S. M., Macheret, V., Sanjuan, J., Drouin, P., & Amarger, N. (2001). Classification of rhizobia based on nodC and nifH gene analysis reveals a close phylogenetic relationship among *Phaseolus vulgaris* symbionts. *Microbiology*, 147(4), 981–993.
- Lakshmanan, V., Selvaraj, G., & Bais, H. (2014). Functional soil microbiome: Belowground solutions to aboveground problem. *Plant Physiology Preview*, 24, pp114.245811.
- Lebeis, L. S. (2014). The potential for give and take in plant–microbiome relationships. *Frontiers in Plant Science | Plant-Microbe Interaction*, 5, 1–6.
- Li, K., Bihan, M., Shibu, Y., & Methé, B. A. (2012). Analyses of the Microbial Diversity across the Human Microbiome. *PLoS ONE* 7(6):, 7(6), e32118.
- Liesack, W., & Dunfield, P. (2004). T-RFLP analysis: A rapid fingerprinting method for studying diversity, structure, and dynamics of microbial communities. In J. Spencer, D. Ragout, & A. Spencer (Eds.), *Environmental Microbiology: Methods and Protocols* (Springer, pp. 23–38). Totowa, New Jersey.
- Liu, C. W., Lin, K. H., & Kuo, Y. M. (2003). Application of factor analysis in the assessment of ground water quality in a black foot disease area in Taiwan. *Science Total Environment*, 313, 77–89.
- Liu, X., Liu, S., Liu, M., Kong, B., Liu, L., & Li, Y. (2014). A primary assessment of the endophytic bacterial community in a xerophilous moss (*Grimmia montana*) using molecular methods and cultivated isolates. *Brazilian Journal of Microbiology*, 45(1), 163–173.
- Liu, Y., Zuo, S., Xu, L., Zou, Y., & Song, W. (2012). Study on diversity of endophytic

- bacterial communities in seeds of hybrid maize and their parental lines. *Archive of Microbiology*, 194(12), 1001–1012.
- Lobato, A. K. D. S., Silveira, J. A. G. Da, Costa, R. C. L. Da, & Neto, C. F. D. O. (2013). Tolerance to Drought in Leguminous Plants Mediated by Rhizobium and Bradyrhizobium. *InTech*. Retrieved from <http://dx.doi.org/10.5772/54094>
- Lopez, B. R., Tinoco-Ojanguren, C., Bacilio, M., Mendoza, A., & Bashan, Y. (2012). Endophytic bacteria of the rock-dwelling cactus *Mammillaria fraileana* affect plant growth and mobilization of elements from rocks. *Environmental and Experimental Botany*, 81, 26–36. <http://doi.org/10.1016/j.envexpbot.2012.02.014>
- López-López, A., Rogel, M. A., Ormeno-Orrillo, E., Martínez-Romero, J., & Martínez-Romero, E. (2010). Phaseolus vulgaris seed-borne endophytic community with novel bacterial species such as Rhizobium endophyticum sp. nov. *Systematic and Applied Microbiology*, 33, 322–327.
- Lopez-Velasco, G., Welbaum, G., Boyer, R., Mane, S., & Ponder, M. (2011). Changes in spinach phylloepiphytic bacteria communities following minimal processing and refrigerated storage described using pyrosequencing of 16S rRNA amplicons. *Journal of Applied Microbiology*, 110(5), 1203–1214.
- Lozupone, C., & Knight, R. (2008). Species divergence and the measurement of microbial diversity. *FEMS Microbiol Review*, 32(4), 557–578.
- Lund, M; Kjeldsen, K; Schramm, A. (2014). The earthworm-Vreminephrobacter symbiosis: an emerging experimental system to study extracellular symbiosis. *Microbial Symbiosis*, 5, 1–6.
- Madeo, D., Comolli, R. L., & Mocenni, C. (2014). Emergence of microbial networks as response to hostile environments. *Frontier in Microbiology*, 5(407), 1–13.
- Makhalanyane, T. P., Valverde, A., Gunnigle, E., Frossard, A., Ramond, J.-B., & Cowan, D. A. (2013). Microbial ecology of hot desert edaphic systems. *FEMS Microbiology Reviews*, 39(2), 203–221. <http://doi.org/10.1093/femsre/fuu011>
- Malfanova, N., Kamilova, F., Validov, S., Shcherbakov, A., Chebotar, V., Tikhonovich, I., & Lugtenberg, B. (2011). Characterization of Bacillus subtilis HC8, a novel plant-beneficial endophytic strain from giant hogweed. *Microbial Biotechnology*, 4, 523–532. <http://doi.org/10.1111/j.1751-7915.2011.00253.x>

- Malik, K. A., & Claus, D. (1979). *Xanthobacter flavus*, a New Species of Nitrogen-Fixing Hydrogen Bacteria. *INTERNATIONAL JOURNAL OF SYSTEMATIC BACTERIOLOGY*, 29(4), 283–287.
- Mandić-Mulec, I., & Prosser, J. (2011). No Title. In E. Logan NA, De Vos P (Ed.), *Endospore-forming soil bacteria, (Soil biology, 27)* (1st ed., pp. 31–59). Springer-Verlag.
- Mandl, K., Schiek, J., Silhavy-Richter, K., Schmidt, H., Schneider, V., & Prange, A. (2014). Vines take up yeast from soil and transport them through the vine to the stem and skin of grapes. *Journal for Ecology, Winegrowing and Climate Faming*, 521, 349–355.
- Mano, H., & Morisaki, H. (2008). Endophytic Bacteria in the Rice Plant. *Microbes and Environments*, 23(2), 109–117. <http://doi.org/10.1264/jsme2.23.109>
- Manter, D., Delgado, J., Holm, D., & Stong, R. (2010). Pyrosequencing reveals a highly diverse and cultivar-specific bacterial endophyte community in potato roots. *Microbial Ecology*, 6(1), 157–166.
- Marasco, R., Rolli, E., Ettoumi, B., Vigani, G., Mapelli, F., Borin, S., ... Daffonchio, D. (2012). A drought resistance-promoting microbiome is selected by root system under desert farming. *PloS One*, 7(10), e48479. <http://doi.org/10.1371/journal.pone.0048479>
- Marasco, R., Rolli, E., Fusi, M., Cherif, A., Abou-Hadid, A., El-Bahairy, U., ... Daffonchio, D. (2013). Plant Growth Promotion Potential Is Equally Represented in Diverse Grapevine Root-Associated Bacterial Communities from Different Biopedoclimatic Environments. *BioMed Research International*, 2013, 1–17.
- Marefat, A., Ophel-Keller, K., & McKay, A. (2004). A real-time PCR assay for detection of *Clavibacter michiganensis* subsp. *insidiosus* in lucerne. *Australasian Plant Pathology*, 36(3), 262–269.
- Marques, J. M., da Silva, T. F., Vollu, R. E., Blank, A. F., Ding, G.-C., Seldin, L., & Smalla, K. (2014). Plant age and genotype affect the bacterial community composition in the tuber rhizosphere of field-grown sweet potato plants. *FEMS Microbiology Ecology*, 88(2), 424–435. <http://doi.org/10.1111/1574-6941.12313>
- Maruatona, G. N., Duodu, K. G., & Minnaar, A. (2010). Physicochemical, nutritional and functional properties of marama bean flour. *Food Chemistry*, 121(2), 400–405. <http://doi.org/10.1016/j.foodchem.2009.12.054>

- Mayali, X., Weber, P. K., Mabery, S., & Pett-Ridge, J. (2014). Phylogenetic Patterns in the Microbial Response to Resource Availability: Amino Acid Incorporation in San Francisco Bay. *PloS One*, 9(4), e95842.
- McInroy, J. a., & Kloepper, J. W. (1995). Survey of indigenous bacterial endophytes from cotton and sweet corn. *Plant and Soil*, 173(2), 337–342.
<http://doi.org/10.1007/BF00011472>
- Mercado-Blanco, J., & Lugtenberg, B. (2014). Biotechnological Applications of Bacterial Endophytes. *Current Biotechnology*, 3, 60–75.
- Minerdi, D., Fani, R., & Gallo, R. (2001). Nitrogen fixation genes in an endosymbiotic Burkholderia strain. *Applied and ...*, 67(2), 725–732.
<http://doi.org/10.1128/AEM.67.2.725>
- Mitchell, R. A. C., Keys, A. J., Madgwick, P. J., Parry, M. A. J., & Lawlor, D. W. (2005). Adaptation of photosynthesis in marama bean *Tylosema esculentum* (Burchell A. Schreiber) to a high temperature, high radiation, drought-prone environment. *Plant Physiology and Biochemistry*, 43(10-11), 969–976.
<http://doi.org/10.1016/j.plaphy.2005.08.009>
- Mitter, B., Brader, G., Afzal, M., Compant, S., Naveed, M., Trognitz, F., & Sessitsch, A. (2013). *Advances in elucidating beneficial interactions between plants, soil, and bacteria. Advances in Agronomy* (Vol. 121). <http://doi.org/10.1016/B978-0-12-407685-3.00007-4>
- Mmonatau, Y. (2005). Flour From the Morama Bean : Composition and Sensory Properties in a Botswana Perspective, (December).
- Montañez, A., Blanco, A. R., Barlocco, C., Beracochea, M., & Sicardi, M. (2012). Characterization of cultivable putative endophytic plant growth promoting bacteria associated with maize cultivars (*Zea mays* L.) and their inoculation effects in vitro. *Applied Soil Ecology*, 58, 21–28. <http://doi.org/10.1016/j.apsoil.2012.02.009>
- Monteiro, M. F., Dutkiewicz, S., & Follows, M. J. (2011). Biogeographical controls on the marine nitrogen fixers. *Global Biochemical Cycles*, 25, GB2003,.
- Morgan, X. C., & Huttenhower, C. (2012). Chapter 12: Human Microbiome Analysis. *PLoS Computational Biology*, 8(12), e1002808.
- Morrissey, J. P., Dow, J. M., Mark, G. L., & O’Gara, F. (2004). Are microbes at the root of a solution to world food production? Rational exploitation of interactions

- between microbes and plants can help to transform agriculture. *EMBO Reports*, 5(10), 922–926. <http://doi.org/10.1038/sj.embor.7400263>
- Mosele, M. M., Hansen, A. S., Engelsens, S. B., Diaz, J., Sørensen, I., Ulvskov, P., ... Harholt, J. (2011). Characterisation of the arabinose-rich carbohydrate composition of immature and mature marama beans (*Tylosema esculentum*). *Phytochemistry*, 72(11-12), 1466–72. <http://doi.org/10.1016/j.phytochem.2011.03.027>
- Muller, H., Berg, C., Landa, B., Auerbach, A., Moissl-Eichinger, C., & Berg, G. (2015). Plant genotype-specific archaeal and bacteria endophytes but similar *Bacillus* antagonists colonize Mediterranean olive trees. *Frontier in Microbiology*, 6, 1–9.
- Muresu, R., Tondello, A., Polone, A., Sulas, L., Baldan, B., & Squartini, A. (2012). Antioxidant treatment counteract the non-culturability of bacteria endophytes isolated from legume nodules. *Archive of Microbiology*, 195(6), 385–391.
- Musler, D., & Schonefeld, H. (2006). The nutrient content of marama bean (*Tylosema esculentum*) an underutilized legume from Southern Africa. *Agricola*, 16, 7–13.
- Muyzer, G. (1999). Genetic fingerprinting of microbial communities – present status and future perspectives. In C. Bell, M. Brylinsky, & P. Johnson-Green (Eds.), *Microbial Biosystems: New Frontiers Proceedings of the 8th International Symposium on Microbial Ecology* (p. 10). Atlantic Canada Society for Microbial Ecology, Halifax, Canada,.
- National Research Council. (2006). *LOST CROPS of AFRICA volume II Vegetables*. Washington DC.
- Naushad, S., Adeolu, M., Goel, N., Khadka, B., Al-Dahwi, A., & Gupta, R. (2015). Phylogenomic and Molecular Demarcation of the Core Members of the Polyphyletic Pasteurellaceae Genera *Actinobacillus*, *Haemophilus*, and *Pasteurella*. *International Journal of Genomics*, 215, 1–14.
- Nei, M., & Kumar, S. (2000). *Molecular Evolution and Phylogenetics*. New York: Press, Oxford University.
- Newton, A. C., Fitt, B. D. L., Atkins, S. D., Walters, D. R., & Daniell, T. J. (2010). Pathogenesis, parasitism and mutualism in the trophic space of microbe-plant interactions. *Trends in Microbiology*, 18(8), 365–73. <http://doi.org/10.1016/j.tim.2010.06.002>
- Newton, A. C., Gravouil, C., & Fountaine, J. M. (2010). Managing the ecology of foliar

- pathogens: Ecological tolerance in crops. *Annals of Applied Biology*, 157(3), 343–359. <http://doi.org/10.1111/j.1744-7348.2010.00437.x>
- Noar, J., & Buckley, D. (2009). *Ideonella azotifigens* sp nov., an aerobic diazotroph of the Betaproteobacteria isolated from grass rhizosphere soil, and emended description of the genus *Ideonella*. *International Journal of Systematic and Evolutionary Microbiology*, 59(8), 1941–1946.
- Nunes, C., Bajji, M., Stepien, V., Mansoa, T., Torres, R., Usall, J., & Jijakli, M. H. (2008). Development and application of a SCAR marker to monitor and quantify populations of the postharvest biocontrol agent *Pantoea agglomerans* CPA-2. *Posthar Biol Technol. Postharvest Biology and Technology*, 47, 422–428.
- Oa, S. B. C. W., Rudrappa, T., Czymmek, K. J., Pare, P. W., Paré, P. W., & Bais, H. P. (2008). Root-Secreted Malic Acid Recruits Beneficial. *Plant Physiology*, 148(November), 1547–1556. <http://doi.org/10.1104/pp.108.127613>
- Ohyama, T., Momose, A., Ohtake, N., Sueyoshi, K., Sato, T., Nakanishi, Y., ... Ando, S. (2014). Nitrogen Fixation in Sugarcane. In T. Ohyama (Ed.), *Advances in Biology and Ecology of Nitrogen Fixation* (pp. 47–70).
- Pacchionni, R., Carvalho, F., Thompson, C., Faustino, A., Nicolini, F., Pereira, T., & Cantao, M. (2014). Taxonomic and functional profiles of soils samples from Atlantic forest and Caatinga biomes in Noertheastern Brazil. *Microbiology Open*, 3(3), 299–315.
- Palmer, C., Bik, E., Eisen, M., Eckburg, P., Sana, T., Wolber, P., ... Brown, P. (2006). Rapid quantitative profiling of complex microbial populations. *Nucleic Acids Research*, 34, e5.
- Parks, D., & Beiko, R. (2012). Measuring community similarity with phylogenetic networks. *Microbiology Evolution*, 29(12), 3947–58.
- Partida-Martínez, L. P., & Heil, M. (2011). The Microbe-Free Plant: Fact or Artifact? *Frontiers in Plant Science*, 2(December), 1–16. <http://doi.org/10.3389/fpls.2011.00100>
- Pedraza, R. O., Mentel, M. I., Ragout, A. L., Xiqui, M. L., Segundo, D. M., & Baca, B. E. (2009). Plant growth-promoting bacteria: the role of chemotaxis in the association *Azospirillum brasilense*-plant. In Frank Columbus (Ed.), *Chemotaxis: Types, Clinical Significance, and Mathematical Models* (pp. 1–31). New York USA: Nova Science Publishers, Inc. USA.

- Per, J. D. H. L., Ryden, A., & Esikuri, E. E. (2005). Capitalizing on the Bio-Economic Value of Multi-Purpose Medicinal Plants for the Rehabilitation of Drylands in Sub-Saharan Africa, 1–53.
- Pereira, P., Ibáñez, F., Rosenblueth, M., Etcheverry, M., & Martínez-Romero, E. (2011). Analysis of the Bacterial Diversity Associated with the Roots of Maize (*Zea mays* L.) through Culture-Dependent and Culture-Independent Methods. *ISRN Ecology*, 2011, 1–10. <http://doi.org/10.5402/2011/938546>
- Peterson, D., Frank, D., Pace, N., & Gordon, J. (2010). Metagenomic approaches for defining the pathogenesis of inflammatory bowel diseases. *Cell Host -Microbes*, 3(6), 417–427.
- Pimentel, I. C., Glienke-Blanco, C., Gabardo, J., Stuart, R. M., & Azevedo, J. L. (2006). Identification and colonization of endophytic fungi from soybean (*Glycine max* (L.) Merrill) under different environmental conditions. *Brazilian Archives of Biology and Technology*, 49(5), 705–711. <http://doi.org/10.1590/S1516-89132006000600003>
- Pisarska, K., & Pietr, S. (2015). Biodiversity of dominant cultivable endophytic bacteria inhabiting tissues of six different cultivars of Maize (*Zea mays* L. ssp *mays*). *Polish Journal of Microbiology*, 64(2), 163–170.
- Poly, F., Monrozier, L., & Bally, R. (2001). Improvement in the RFLP procedure for studying the diversity of *nifH* genes in communities of nitrogen fixers in soil. *Research in Microbiology*, 152(1), 95–103.
- Poppe, L., Vanhoutte, S., & Hofte, M. (2003). Modes of action of *Pantoea agglomerans* CPS-2, an antagonist of postharvest pathogens on fruits. *European Journal of Plant Pathology*, 109, 963–973.
- Pradeu, T. (2011). A Mixed Self: The Role of Symbiosis in Development. *Biological Theory*, 6(1), 80–88.
- Puspasari, I., Talib, M. Z. M., Daud, W. R. W., & Tasirin, S. M. (2014). Characteristic Drying Curve of Oil Palm Fibers. *International Journal on Advance Science Engineering Information Technology*, 4(1), 20–24.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., ... Glöckner, F. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, 2013, Vol. 41, 41, D590–D596.

- Raaijmakers, J. M., Paulitz, T. C., Steinberg, C., Alabouvette, C., & Moëgne-Loccoz, Y. (2009). The rhizosphere: a playground and battlefield for soilborne pathogens and beneficial microorganisms. *Plant and Soil*, 321(1-2), 341–361.
<http://doi.org/10.1007/s11104-008-9568-6>
- Ramírez-Puebla, S. T., Servín-Garcidueñas, L. E., Jiménez-Marín, B., Bolaños, L. M., Rosenblueth, M., Martínez, J., ... Martínez-Romero, E. (2013). Gut and root microbiota commonalities. *Applied and Environmental Microbiology*, 79(1), 2–9.
<http://doi.org/10.1128/AEM.02553-12>
- Ramos, H. J. O., Yates, M. G., Pedrosa, F. O., & Souza, E. M. (2011). Strategies for bacterial tagging and gene expression in plant-host colonization studies. *Soil Biology and Biochemistry*, 43(8), 1626–1638.
<http://doi.org/10.1016/j.soilbio.2011.03.029>
- Rastogi, G., & Sani, R. K. (2011). *Microbes and Microbial Technology*, 29–58.
<http://doi.org/10.1007/978-1-4419-7931-5>
- Raymon, S., Esther, H., Ramble, A., Abasiofiok M., I., Ronald C, S., & Scot, D. (2013). Assessing the Diversity and Composition of Bacterial Communities across a Wetland, Transition, Upland Gradient in Macon County Alabama. *Diversity*, 5(3), 461–478.
- Redman, R. S., Kim, Y. O., Woodward, C. J. D. A., Greer, C., Espino, L., Doty, S. L., & Rodriguez, R. J. (2011). Increased Fitness of Rice Plants to Abiotic Stress Via Habitat Adapted Symbiosis: A Strategy for Mitigating Impacts of Climate Change. *PLoS ONE*, 6(7), e14823. <http://doi.org/10.1371/journal.pone.0014823>
- Reinhold, B., Hurek, T., Niemann, E. G., & Fendrik, I. (1986). Close association of *Azospirillum* and diazotrophic rods with different root zones of Kallar grass. *Applied and Environmental Microbiology*, 52(3), 520–526.
- Reinhold-Hurek, B., & Hurek, T. (2000). Reassessment of the taxonomic structure of the diazotrophic genus *Azoarcus* sensu lato and description of three new genera and new species, *Azovibrio restrictus* gen. nov., sp. nov., *Azospira oryzae* gen. nov., sp. nov. and *Azonexus fungiphilus* gen. nov., s. *International Journal of Systematic and Evolutionary Microbiology*, 50, 649–659.
- Reinhold-Hurek, B., & Hurek, T. (2011). Living inside plants: bacterial endophytes. *Current Opinion in Plant Biology*, 14(4), 435–443.
<http://doi.org/10.1016/j.pbi.2011.04.004>

- Reisberg, E. E., Hildebrandt, U., Riederer, M., & Hentschel, U. (2013). Distinct phyllosphere bacterial communities on *Arabidopsis wax* mutant leaves. *PloS One*, 8(11), e78613. <http://doi.org/10.1371/journal.pone.0078613>
- Reiter, B., & Sessitsch, A. (2006). Bacterial endophytes of the wildflower *Crocus albiflorus* analyzed by characterization of isolates and by a cultivation-independent approach. *Canadian Journal of Microbiology*, 52(2), 140–149.
- Rey, T., & Schornack, S. (2013). Interactions of beneficial and detrimental root-colonizing filamentous microbes with plant hosts. *Genome Biology*, 14(6), 121. <http://doi.org/10.1186/gb-2013-14-6-121>
- Rezzonico, F., Smits, T., & Duffy, B. (2012). Detection of AI-2 Receptors in Genomes of Enterobacteriaceae Suggests a Role of Type-2 Quorum Sensing in Closed Ecosystems. *Sensors*, 12, 6645–6665.
- Robinson, C.J; Bohannan; B.J; Young, V. (2010). From structure to function: the ecology of host-associated microbial communities. *Microbiology and Molecular Biology Reviews : MMBR*, 74(3), 453–76.
- Rodrigues ML, Nakayasu ES, Oliveira DL, Nimrichter L, Nosanchuk JD, Almeida IC, C. A. (2008). Extracellular vesicles produced by *Cryptococcus neoformans* contain protein components associated with virulence. *Eukaryotic Cell*, 7(1), 58–67.
- Rosch, C., Mergel, A., & Bothe, H. (2002). Biodiversity of denitrifying and dinitrogen-fixing bacteria in an acid forest soil. *Applied and Environmental Microbiology*, 68, 3818–3829.
- Rosenberg, E., & Zilber-Rosenberg, I. (2011). Symbiosis and development: The hologenome concept. *Birth Defects Research Part C: Embryo Today: Reviews*, 93(1), 56–66. <http://doi.org/10.1002/bdrc.20196>
- Rosenblueth, M., & Martínez-Romero, E. (2006). Bacterial endophytes and their interactions with hosts. *Molecular Plant-Microbe Interactions : MPMI*, 19(8), 827–837. <http://doi.org/10.1094/MPMI-19-0827>
- Rout, M. E., & Southworth, D. (2013). The root microbiome influences scales from molecules to ecosystems: The unseen majority. *American Journal of Botany*, 100(9), 1689–1691. <http://doi.org/10.3732/ajb.1300291>
- Ruppel, S. C., Hecht-Buchholz, R., Remus, U., Ortmann, & Schmelzer, R. (1992). Settlement of the diazotrophic, phytoeffective bacterial strain *Pantoea agglomerans*

- on winter wheat: an investigation using ELISA and transmission electron microscopy. . *Plant Soil*, 145, 261–273.
- Russo, A., Carrozza, G. Pietro, Vettori, L., Felici, C., Cinelli, F., & Toffanin, A. (2012). Plant Beneficial Microbes and Their Application in Plant Biotechnology . *Innovations in Biotechnology*. <http://doi.org/10.5772/2450>
- Sagaram, U. S., Deangelis, K. M., Trivedi, P., Andersen, G. L., Lu, S. E., & Wang, N. (2009). Bacterial diversity analysis of huanglongbing pathogen-infected citrus, using phyloChip arrays and 16S rRNA gene clone library sequencing. *Applied and Environmental Microbiology*, 75(6), 1566–1574. <http://doi.org/10.1128/AEM.02404-08>
- Saharan, B. S., & Nehra, V. (2011). Plant Growth Promoting Rhizobacteria : A Critical Review. *Life Sciences and Medicine Research*, 2011, 1–30.
- Santhanam, R., Groten, K., Meldau, D. G., & Baldwin, I. T. (2014). Analysis of plant-bacteria interactions in their native habitat: Bacterial communities associated with wild tobacco are independent of endogenous jasmonic acid levels and developmental stages. *PLoS ONE*, 9(4), 1–12. <http://doi.org/10.1371/journal.pone.0094710>
- Sato, A., Watanabe, T., Unno, Y., Purnomo, E., Osaki, M., & Shinano, T. (2009). Analysis of Diversity of Diazotrophic Bacteria Associated with the Rhizosphere of a Tropical Arbor, *Melastoma malabathricum* L. *Microbes and Environments*, 24(2), 81–87. <http://doi.org/10.1264/jsme2.ME08565>
- Saunders, M., Glenn, A. E., & Kohn, L. M. (2010). Exploring the evolutionary ecology of fungal endophytes in agricultural systems: using functional traits to reveal mechanisms in community processes. *Evolutionary Applications*, 3(5-6), 525–537.
- Schenk, P. M., Carvalhais, L. C., & Kazan, K. (2012). Unraveling plant–microbe interactions: can multi-species transcriptomics help? *Trends in Biotechnology*, 30(3), 177–184. <http://doi.org/10.1016/j.tibtech.2011.11.002>
- Schmit, J. P., & Lodge, D. J. (2005). Classical Methods and Modern Analysis for Studying Fungal Diversity. In J. Dighton, J. F. White Jr., & P. Oudemans (Eds.), *The Fungal Community: Its Organization and Role in the Ecosystem* (Third, p. 960). CRC Press.
- Schoss, P. D., & Handelsman, J. (2004). No Title. *Microbiology and Molecular Review*, 68(4), 686–691.

- Schulz, B., & Boyle, C. (2006). What are Endophytes? In B. Schulz & C. Boyle (Eds.), *Microbial root endophytes* (Vol. 9, pp. 1–14). Berlin, Heidelberg: Springer.
- Scott, F. G., Sapp, J., & Tauber, A. I. (2012). A SYMBIOTIC VIEW OF LIFE: WE HAVE NEVER BEEN INDIVIDUALS. *The Quarterly Review of Biology*, 87(4), 325–341.
- Seo, W., Lim, W., Kim, E., Yun, H., & Cho, K. (2010). Endophytic bacterial diversity in the young radish and their antimicrobial activity against pathogens. *Journal of the Korean Society for Applied Biological Chemistry*, 53(4), 493–503.
- Sessitsch, A., Reiter, B., & Berg, G. (2004). Endophytic bacterial communities of field-grown potato plants and their plant-growth-promoting and antagonistic abilities. *Canadian Journal of Microbiology*, 50(4), 239–249. <http://doi.org/10.1139/w03-118>
- Sessitsch, Hardoim, P., Döring, J., Weilharter, A., Krause, A., Woyke, T., ... Reinhold-Hurek, B. (2012). Functional Characteristics of an Endophyte Community Colonizing Rice Roots as Revealed by Metagenomic Analysis. *Molecular Plant-Microbe Interactions*, 25(1), 28–36. <http://doi.org/10.1094/MPMI-08-11-0204>
- Shelef, O., Helman, Y., Friedman, A.-L.-L., Behar, A., & Rachmilevitch, S. (2013). Tri-Party Underground Symbiosis between a Weevil, Bacteria and a Desert Plant. *PLoS ONE*, 8(11), e76588. <http://doi.org/10.1371/journal.pone.0076588>
- Shia, S., Nuccioa, E., Hermana, D., Rijkersa, R., Esteraa, K., Lib, J., ... Firestonea, M. (2015). Successional Trajectories of Rhizosphere Bacterial Communities over Consecutive Seasons. *mBio*, 6(4), e00746–15.
- Simonet, P., Grosjean, M., Misra, A., Nazaret, S., Cournoyer, B., & Normand, P. (1991). Frankia genus-specific characterization by polymerase chain reaction. *Applied Environmental Microbiology*, 57, 3278–3286.
- Singh, S., Pandey, A., Kumar, B., & Palni, L. M. S. (2010). Enhancement in growth and quality parameters of tea [*Camellia sinensis* (L.) O. Kuntze] through inoculation with arbuscular mycorrhizal fungi in an acid soil. *Biology and Fertility of Soils*, 46(5), 427–433.
- Siqueira, J. J., ; Sakamoto, M., & Rosado, A. (2010). Microbial community profiling using terminal restriction fragment length polymorphism (T-RFLP) and denaturing gradient gel electrophoresis (DGGE). *Methods in Molecular Biology*, 666, 71–85.

- Steenhoudt, O., & Vanderleyden, J. (2000). Azospirillum, a free-living nitrogen-fixing bacterium closely associated with grasses: genetic, biochemical and ecological aspects. *FEMS Microbiology Reviews*, 24(4), 487–506.
- Stratil, S., Neulinger, S., Knecht, H., Friedrichs, A., & Wahl, M. (2013). Temperature-driven shifts in the epibiotic bacterial community composition of the brown macroalga *Fucus vesiculosus*. *Microbiology Open*, 2(2), 338–349.
- Štursa, P., Uhlík, O., Kurzawová, V., Koubek, J., Ionescu, M., Strohalm, M., ... Macková, M. (2009). No Title. *Plant, Soil and Environment*, 55, 389–396.
- Sturz, A. V., Christie, B. R., & Matheson, B. G. (1998). Associations of bacterial endophyte populations from red clover and potato crops with potential for beneficial allelopathy. *Canadian Journal of Microbiology*, 44, 162–167.
- Su, C., Lei, L., Duan, Y., Zhang, K., & Yang, J. (2012). Culture-independent methods for studying environmental microorganisms: methods, application, and perspective. *Appl Microbiol Biotechnol*. 2012 Feb;93(3):993-1003, 93(3), 993–100.
- Sun, L., Qiu, F., Zhang, X., Dai, X., Dong, X., & Song, W. (2008). Endophytic Bacterial Diversity in Rice (*Oryza sativa* L.) Roots Estimated by 16S rDNA Sequence Analysis. *Microbial Ecology*, 55(3), 415–424. <http://doi.org/10.1007/s00248-007-9287-1>
- Taghavi, S., & Lelie, D. van der. (2013). Genome Sequence of the Plant Growth-Promoting Endophytic Bacterium *Enterobacter* sp. 638. In F. J. de Bruijn (Ed.), *Molecular Microbial Ecology of the Rhizosphere: Volume 1 & 2* (p. 1328). Wiley-Blackwell.
- Taghavi, S., van der Lelie, D., Hoffman, A., Zhang, Y.-B., Walla, M. D., Vangronsveld, J., ... Monchy, S. (2010). Genome Sequence of the Plant Growth Promoting Endophytic Bacterium *Enterobacter* sp. 638. *PLoS Genetics*, 6(5), e1000943. <http://doi.org/10.1371/journal.pgen.1000943>
- 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.
- Tikhonovich, I. A., & Provorov, N. A. (2011). Microbiology is the basis of sustainable agriculture: an opinion. *Annals of Applied Biology*, 159(2), 155–168. <http://doi.org/10.1111/j.1744-7348.2011.00489.x>

- Tiquia, S. M. (2010). Using Terminal Restriction Fragment Length Polymorphism (T-RFLP) Analysis to Assess Microbial Community Structure in Compost Systems. *Methods in Molecular Biology*, 599, 89–102.
- Tompson, J., Gibson, T., Plewniak, F., Jeanmougin, F., & Higgins, D. (1997). The CLUSTAL-X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acid Research*, 25, 4876–4882.
- Torres, A., Araújo, W., Cursino, L., Hungria, M., Plotegher, F., Mostasso, F., & Azevedo, J. (2008). Diversity of endophytic enterobacteria associated with different host plants. *Journal of Microbiology*, 46(4), 373–379.
- Trabelsi, D., & Mhamdi, R. (2013). Microbial Inoculants and Their Impact on Soil Microbial Communities: A Review. *BioMed Research International*, 2013, 1–11.
- Travlos, I. S., Economou, G., & Karamanos, a. I. (2007). Germination and emergence of the hard seed coated *Tylosema esculentum* (Burch) A. Schreib in response to different pre-sowing seed treatments. *Journal of Arid Environments*, 68(3), 501–507. <http://doi.org/10.1016/j.jaridenv.2006.07.001>
- Travlos, I. S., Liakopoulos, G., Karabourniotis, G., Fasseas, C., & Karamanos, a. J. (2008). Circadian leaflet movements of *Tylosema esculentum* (Burch) A. Schreib, and the abolishment of these diurnal movements by potassium deficiency. *Journal of Arid Environments*, 72(9), 1745–1750. <http://doi.org/10.1016/j.jaridenv.2008.03.020>
- Trotel-Aziz, P., Couderchet, M., Biagiante, S., & Aziz Aziz, A. (2008). Characterization of new bacterial biocontrol agents *Acinetobacter*, *Bacillus*, *Pantoea* and *Pseudomonas* spp. mediating grapevine resistance against *Botrytis cinerea*. *Environmental and Experimental Botany*, 64, 21–32.
- Truyens, S., Weyens, N., Cuypers, A., & Vangronsveld, J. (2015). Bacterial seed endophytes: genera, vertical transmission and interaction with plants. *Environmental Microbiology Reports*, 7(1), 40–50.
- Tsavkelova, E., Cherdyntseva, T., Klimova, S., Shestakov, A., Botina, S., & Netrusov, A. (2007). Orchid-associated bacteria produce indole-3-acetic acid, promote seed germination, and increase their microbial yield in response to exogenous auxin. *Archive of Microbiology*, 188(6), 188(6):655–664.
- Tsiamis, G., Karpouzas, D., Cherif, A., & Mavrommatis, K. (2014). Microbial Diversity for Biotechnology. In G. Tsiamis, D. Karpouzas, A. Cherif, & K. Mavrommatis

- (Eds.), *Microbial diversity for Biotechnology* (p. 305).
- Tsiamis, S., Nikolaki, & George. (2013). Microbial Diversity in the Era of Omic Technologies. *BioMed Research International*, 2013 (2013, 1–15.
- Ulrich, K., Ulrich, A., & Ewald, D. (2008). Diversity of endophytic bacterial communities in poplar grown under field conditions. *FEMS Microbiology Ecology*, 63(2), 169–180. <http://doi.org/10.1111/j.1574-6941.2007.00419.x>
- Vacheron, J., Desbrosses, G., Bouffaud, M.-L., Touraine, B., Moëgne-Loccoz, Y., Muller, D., ... Prigent-Combaret, C. (2013). Plant growth-promoting rhizobacteria and root system functioning. *Frontiers in Plant Science*, 4(September), 356. <http://doi.org/10.3389/fpls.2013.00356>
- Valášková, V., & Baldrian, P. (2009). Denaturing gradient gel electrophoresis as a fingerprinting method for the analysis of soil microbial communities. *Plant Soil and Environment* 55 (2009): 413–423, 55, 413–423.
- Vautrin E, V. F. (2009). Interactions between vertically transmitted symbionts: cooperation or conflict? *Trends in Microbiology*, 17(3), 95–99.
- Verstraete, B., Janssens, S., Lemaire, B., Smets, E., & Dessein, S. (2013). Phylogenetic lineages in Vanguerieae (Rubiaceae) associated with Burkholderia bacteria in sub-Saharan Africa. *American Journal of Botany*, 100(12), 2380–7. <http://doi.org/10.3732/ajb.1300303>
- Vetrivelkalai, P., Sivakumar, M., & Jonathan, E. I. (2010). Biocontrol potential of endophytic bacteria on Meloidogyne incognita and its effect on plant growth in bhendi. *Journal of Biopesticides*, 3(2), 452–457.
- Větrovský, T., & Baldrian, P. (2013). The Variability of the 16S rRNA Gene in Bacterial Genomes and Its Consequences for Bacterial Community Analyses. *PLoS ONE*, 8(2), e57923.
- Wang, L., D’Odorico, P., Ringrose, S., Coetzee, S., & Macko, S. (2007). Biogeochemistry of the Kalahari sands. *Journal of Arid Environments*, 71, 259–279.
- Wang, X., Cai, Y., Sun, Y., Knight, R., & Mai, V. (2012). Secondary structure information does not improve OTU assignment for partial 16s rRNA sequences. *The ISME Journal* (2012) 6, 6, 1277–1280.
- Wang, W; Vinocur, B; Altman, A. (2003). Plant responses to drought, salinity and

- extreme temperatures: towards genetic engineering for stress tolerance. *Planta*, 218, 1–14.
- Wehner, J. (2013). *Interactions between root-associated fungi and their host plants in a dry grassland*. Freie Universität Berlin.
- Weisburg, W., Barns, S., Pelletier, D., & Lane, D. (1991). 16S ribosomal DNA amplification for phylogenetic study. *Journal of Bacteriology*, 173, 697–703.
- White, J. R., Patel, J., Ottesen, A., Arce, G., Blackwelder, P., & Lopez, J. V. (2012). Pyrosequencing of Bacterial Symbionts within *Axinella corrugata* Sponges: Diversity and Seasonal Variability. *PLoS ONE*, 7(6), e38204.
- Wong, M., & Medrano, J. (2005). Real-time PCR for mRNA quantitation. *BioTechniques*, 39, 75–85.
- Yoshida, S. (1981). Physiological analysis of rice yield. In International Rice Research Institute (Ed.), *Fundamentals of rice crop science* (pp. 231–251). Makita: Makita City (Philippines).
- Yule, K. M., Miller, T. E. X., & Rudgers, J. A. (2013). Costs, benefits, and loss of vertically transmitted symbionts affect host population dynamics. *Oikos*, (December 2012), no–no. <http://doi.org/10.1111/j.1600-0706.2012.00229.x>
- Yu-Qing, C., Guan-Ju, R., Shu-Qing, A., Qing-Ye, S., LChang-Hong, L., & Shuang, J.-L. (2008). Changes of Bacterial Community Structure in Copper Mine Tailings After Colonization of Reed (*Phragmites communis*). *Pedosphere*, 18(6), 731–740. [http://doi.org/10.1016/S1002-0160\(08\)60068-5](http://doi.org/10.1016/S1002-0160(08)60068-5)
- Zawonzik, M., Vazquez, S., Herrera, S., & Groppa, M. (2014). Search for endophytic diazotroph in barley seeds. *Brazilian Journal of Microbiology*, 45(2), 621–625.
- Zeng, Y., Feng, F., Medová, H., Dean, J., & Koblížek, M. (2014). Functional type 2 photosynthetic reaction centers found in the rare bacterial phylum Gemmatimonadetes. *Proceedings of the National Academy of Sciences*, 111(21), 7795–7800.
- Zgadzaj, R., James, E. K., Kelly, S., Kawaharada, Y., de Jonge, N., Jensen, D. B., ... Radutoiu, S. (2015). A Legume Genetic Framework Controls Infection of Nodules by Symbiotic and Endophytic Bacteria. *PLOS Genetics*, 11(6), e1005280. <http://doi.org/10.1371/journal.pgen.1005280>
- Zhan, J., & Sun, Q. (2011). Diversity of free-living nitrogen-fixing microorganisms in

wastelands of copper mine tailings during the process of natural ecological restoration. *Journal of Environmental Sciences (China)*, 23(3), 476–487.

Zhang, H., John, R., Peng, Z., Yuan, J., Chu, C., Du, G., & Zhou, S. (2012). The relationship between species richness and evenness in plant communities along successional gradient: A study from Sub-Alpine Meadow of the Eastern Qinghai-Tibetan Plateau. *PloS One*, 7(11), e49024.

Zhong, K., Gao, X.-L., Xu, Z.-J., Gao, H., Fan, S., Yamaguchi, I., ... Chen, R.-J. (2011). Antioxidant Activity of a Novel *Streptomyces* Strain Eri12 Isolated from the Rhizosphere of *Rhizoma curcuma* Longae. *Current Research in Bacteriology*, 4(2), 63–72. <http://doi.org/10.3923/crb.2011.63.72>

Zhou, T., Yang, L., Lu, Y., Dror, I., Dantas Machado, A. C., Ghane, T., ... Rohs, R. (2013). DNASHape: a method for the high-throughput prediction of DNA structural features on a genomic scale. *Nucleic Acids Research*, 41(Web Server issue), W56–62. <http://doi.org/10.1093/nar/gkt437>

APPENDIX

Appendix 1: Diversity of culturable bacteria within different samples

Sample	<i>Streptococcus</i>	<i>Lactococcus</i>	<i>Brevibacillus</i>	<i>Lysinibacillus</i>	<i>Pantibacillus</i>	Unclassified <i>Bacillus</i>	Unclassified <i>Arthrobacter</i>	<i>Curtobacterium</i>	Unclassified <i>Streptomyces</i>	Actinomycetales	<i>Ochrobactrum</i>	<i>Rhizobium</i>	<i>Achromobacter</i>	<i>Burkholderia</i>	<i>Acinetobacter</i>	<i>Azomonas</i>	<i>Cosmetia</i>	<i>Enterobacter</i>	<i>Escherichia/Shigella</i>	<i>Klebsiella</i>	<i>Kosakonia</i>	<i>Pantoea</i>	<i>Stenotrophomonas</i>	Unclassified <i>Pseudomonas</i>	Unclassified <i>Enterobacteriaceae</i>	Unclassified <i>Trabulsiella</i>	Unclassified <i>Pseudomonadaceae</i>	Unclassified <i>Xanthomonadaceae</i>
LOM2014	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	3	0	0	0	0	0
LOM2013	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	1	0	0	2	0	1	0	0	0	0	0	2	0
LOM2012	0	0	0	0	0	3	0	0	0	2	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	2	0
LOM2011	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	3	0	0	1	0	0	0	1	2	1	0	2	0
LHARN2014	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	2	0	0	0	2	0
LHARN2013	0	0	0	1	0	2	1	0	0	0	0	0	0	0	0	1	0	0	1	0	1	0	0	1	0	0	1	0
LHARN2012	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0
LHARN2011	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	1	0	0	2	0	1	0	0	1	1	0	4	0
LOTJ2014	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	2	0	0	2	3	0	0	0	1	0
LOTJ2013	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	1	0	0	2	0	0	0	0	0	0	0	1	0
LOTJ2012	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	2	0	0	1	1	0	0	0	4	0
LOTJ2011	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	2	0	0	1	0	0	0	0	1	1	0	2	0
SOM2014	0	0	1	0	0	1	1	0	0	2	1	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	7	0
SOM2013	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	2	0	0	1	0	1	0	0	2	0	0	1	0
SOM2012	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0
SOM2011	0	0	0	0	0	3	0	1	0	0	0	0	0	0	0	2	0	0	2	0	0	0	0	1	0	0	3	0
SHARN2014	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
SHARN2013	0	0	0	2	0	1	0	1	0	0	0	0	0	0	0	1	0	0	2	0	0	0	0	1	0	0	1	0
SHARN2012	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0	0	3	0
SHARN2011	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	2	0	0	2	0	0	0	0	1	0	0	3	0

		Unclassified Xanthomonadaceae																												
		Unclassified Pseudomonadaceae																												
		Unclassified Enterobacteriaceae																												
		Unclassified Tracheibacteriaceae																												
		Unclassified Pseudomonas																												
		Unclassified Pantoea																												
		Unclassified Klebsiella																												
		Unclassified Escherichia/Shigella																												
		Unclassified Enterobacter																												
		Unclassified Coxiella																												
		Unclassified Azomonas																												
		Unclassified Acinetobacter																												
		Unclassified Burkholderia																												
		Unclassified Achromobacter																												
		Unclassified Rhizobium																												
		Unclassified Ochrobactrum																												
		Unclassified Actinomycetales																												
		Unclassified Streptomyces																												
		Unclassified Curtobacterium																												
		Unclassified Bacillus																												
		Unclassified Bacilli																												
		Unclassified Paenibacillus																												
		Unclassified Lysinibacillus																												
		Unclassified Brevibacillus																												
		Unclassified Lactococcus																												
		Unclassified Streptococcus																												
	Sample																													
	SOTJ2014	0	0	0	3	0	1	1	0	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	0	0	1	0	1	
	SOTJ2013	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	2	0	0	0	0	1	0	0	2	0	0	
	SOTJ2012	0	0	0	0	0	3	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	2	0	0	0	2	0	0
	SOTJ2011	0	0	0	0	0	4	0	0	0	0	0	0	2	0	0	2	0	0	3	0	0	0	0	1	1	0	3	0	0
	TOM2014	0	0	0	2	0	1	0	0	0	0	0	0	1	0	0	1	0	0	6	1	2	0	0	1	0	0	0	0	0
	TOM2013	1	0	0	0	0	9	0	0	0	0	0	0	1	0	1	8	0	0	3	1	1	0	2	1	0	0	6	0	1
	TOM2012	0	0	1	0	0	6	1	0	0	0	0	0	1	0	1	2	0	0	3	1	2	1	2	0	1	0	8	0	1
	TOM2011	0	0	1	1	0	9	1	1	1	2	1	1	0	0	0	3	1	0	4	1	2	1	1	2	2	1	7	0	1
	THARN2014	0	1	0	2	0	1	0	0	0	1	0	0	0	0	0	1	0	1	4	0	2	0	0	1	0	0	6	1	0
	THARN2013	0	1	0	0	0	13	0	0	0	0	0	0	1	0	0	7	0	0	3	1	1	0	2	1	0	0	9	0	0
	THARN2012	0	0	1	0	1	8	0	0	0	1	0	0	0	0	0	1	0	0	5	0	4	0	5	0	0	0	9	0	0
	THARN2011	0	0	0	1	0	7	0	1	0	2	0	1	0	1	0	4	0	0	3	0	2	0	1	2	2	0	3	0	0
	TOTJ2014	1	0	0	0	0	1	0	0	0	2	0	0	0	0	0	0	0	0	9	1	2	1	0	1	0	0	2	0	0
	TOTJ2013	0	0	0	3	0	16	0	0	0	0	0	0	0	0	0	6	0	0	3	0	1	0	0	3	1	0	11	0	0
	TOTJ2012	0	0	0	1	0	7	0	0	0	0	0	0	0	0	0	3	0	0	4	0	1	0	2	0	0	0	10	0	0
	TOTJ2011	0	0	0	0	0	8	0	0	0	1	0	0	0	0	0	3	0	0	4	0	2	0	2	1	2	0	6	0	0

APPENDIX 2. TABLES CHAPTER FOUR

Table 10: Sequence similarity between diazotrophic bacterial isolates and the closest type strain from the NCBI Database. A Comparison between the first hit and cultured bacteria.

Taxonomic group and strain		Nearest species relative (Cultured)	% Identity nucleotide	Nearest species relative (First hit)	% Identity nucleotide
Alphaproteobacteria	6_AQER	<i>Xanthobacter flavus</i>	92	Uncultured marine bacterium clone: KS-14-2_ON1_16.	99
Alphaproteobacteria	9_AQER	<i>Azospirillum brasilense</i>	99	Uncultured bacterium clone L21-H146	100
Alphaproteobacteria	11_AQER	<i>Azospirillum zeae</i>	99	<i>Azospirillum zeae</i> strain Gr31	99
Alphaproteobacteria	13_AQER	<i>Xanthobacter autotrophicus</i>	93	Uncultured marine bacterium clone: KS-14-2_ON1_16.	99
Alphaproteobacteria	14_AQER	<i>Xanthobacter autotrophicus</i>	87	Uncultured bacterium clone GR-F6	93
Alphaproteobacteria	UZ15_AQER	<i>Methylosinus</i> sp. LW4	85	Uncultured bacterium clone L21-H187	93
Alphaproteobacteria	22_AQER	<i>Rhodovulum adriaticum</i>	90	<i>Rhodovulum adriaticum</i> strain 2781	90
Alphaproteobacteria	25_AQER	<i>Azospirillum brasilense</i>	89	Uncultured bacterium clone SIW1-2	90
Alphaproteobacteria	26_AQER	<i>Magnetospirillum magneticum</i>	88	Uncultured bacterium clone IFRpool-20	89
Alphaproteobacteria	28_AQER	<i>Azospirillum brasilense</i>	88	Uncultured bacterium clone IFRpool-20	89
Alphaproteobacteria	29_AQER	<i>Xanthobacter autotrophicus</i>	92	Uncultured marine bacterium	96
Alphaproteobacteria	31_AQER	<i>Xanthobacter autotrophicus</i>	92	Uncultured marine bacterium	99
Alphaproteobacteria	52_AQER	<i>Rhodovulum euryhalinum</i>	90	Uncultured bacterium clone IFRpool-47	92
Alphaproteobacteria	45_AQER	<i>Rhodovulum</i> sp. CP-10	90	<i>Rhodovulum</i> sp. CP-10	90
Alphaproteobacteria	56_AQER	<i>Xanthobacter flavus</i>	94	Uncultured marine bacterium clone: KS-14-2_ON1_16	99
Alphaproteobacteria	49_AQER	<i>Xanthobacter autotrophicus</i>	87	Uncultured bacterium clone GR-F6	93
Alphaproteobacteria	57_AQER	<i>Rhodovulum</i> sp. CP-10	90	Uncultured bacterium clone 1-08	91
Alphaproteobacteria	59_AQER	<i>Rhodovulum sulfidophilum</i>	90	Uncultured bacterium clone 1-08	91

Table 11: Sequence similarity between diazotrophic bacterial isolates and the closest type strain from the NCBI Database. A Comparison between the first hit and cultured bacteria (Continued).

Bacillales	17 A QER	<i>Bacillus</i> sp. ZN20	99	<i>Bacillus</i> sp. ZN20		99
Bacillales	27 A QER	<i>Bacillus</i> sp. ZN20	99	<i>Bacillus</i> sp. ZN20		99
Bacillales	36 A QER	<i>Paenibacillus typhos</i>	94	<i>Paenibacillus</i> sp. xj7		94
Betaproteobacteria	5B A QER	<i>Burkholderia fungorum</i>	91	Uncultured bacterium clone I21-H187		95
Betaproteobacteria	8B A QER	<i>Burkholderia cepacia</i>	87	Uncultured bacterium clone I21-H187		88
Betaproteobacteria	12 A QER	<i>Burkholderia silvatlantica</i>	99	<i>Burkholderia silvatlantica</i> strain SRCL318		99
Betaproteobacteria	UZ 9	<i>Burkholderia fungorum</i>	92	Uncultured bacterium clone I21-H187		93
Betaproteobacteria	33 A QER	<i>Burkholderia silvatlantica</i>	98	<i>Burkholderia silvatlantica</i> strain SRCL318		98
Betaproteobacteria	37 A QER	<i>Rubrivivax gelatinosus</i>	88	Uncultured bacterium clone 08826-25 (MOTU1539)		92
Betaproteobacteria	39 A QER	<i>Burkholderia cepacia</i>	98	Uncultured microorganism clone JUL_H04_RNA_C07		99
Betaproteobacteria	41 A QER	<i>Idoneella dechloratans</i>	93	Uncultured bacterium clone I836		94
Betaproteobacteria	46 A QER	<i>Idoneella dechloratans</i>	91	Uncultured bacterium clone MN24		92
Betaproteobacteria	54 A QER	<i>Burkholderia cepacia</i>	99	Uncultured microorganism clone H05_RNA_G08		99
Betaproteobacteria	58 A QER	<i>Pelomonas saccharophila</i>	91	Uncultured marine bacterium clone: KS-14-2_ON1_16		95
Betaproteobacteria	NN2 A QER	<i>Burkholderia fungorum</i>	92	Uncultured bacterium clone I21-H187		96
Betaproteobacteria	UZ3 A QER	<i>Burkholderia fungorum</i>	93	Uncultured bacterium clone I21-H187		95
Betaproteobacteria	UZ9 A QER	<i>Burkholderia fungorum</i>	92	Uncultured bacterium clone I21-H187		93
Gamma proteobacteria	3B A QER	<i>Enterobacter</i> sp.	87	Uncultured bacterium clone IPA30		85
Gamma proteobacteria	UZ1 A QER	<i>Stenotrophomonas</i>	97	<i>Stenotrophomonas</i> sp. isolate gx-44		97
Oscillatoriophyceae	18 A QER	<i>Plectonema cf. radiosum</i>	95	<i>Plectonema cf. radiosum</i> LEGE 06105		95
	10B A QER	<i>Bacterium DZT-HS14</i>	96	<i>Bacterium DZT-HS14</i>		96
Betaproteobacteria	4B A QER	<i>Burkholderia fungorum</i>	92	Uncultured bacterium clone I21-H187		96

Table 12: Eigenvalues and percentage of explained variability extracted from *T. esclentum* sequences based on 16S rRNA gene sequences generated using Roche/454 pyrosequencing.

PC	Eigenevalue	%Variance
1	223343	94.00
2	8938.60	3.80
3	4001.40	2.00
4	893.68	0.40
5	276.40	0.12
6	96.80	0.04
7	20.39	0.01
8	7.30	0.003
9	1.73E-28	7.28E-32

APPENDIX 3: Change in diversity of total endophytic microbial communities in leaves, stems and tuber of *T.esculentum* over time and location. OM, HARN, OTJ represent Omirata, Harnas and Otjinene respectively while the number represent the collection year.

Sample	Shannon's Index (H)L	Shannon's index (H) S	Shannon's Index (H)T		Simpson_1-DL	Simpson_1-DS	Simpson_1-DT	Evenness_e^H/SL	Evenness_e^H/Ss	Evenness_e^H/ST	Chao-1L	Chao-1S
OM2014	1.475	1.708	1.807	0.7347	0.7378	0.7822	0.8743	0.6894	0.7611	11	15.5	11.33
OM2013	1.889	1.889	2.074	0.8395	0.8395	0.8359	0.9448	0.9448	0.6631	10.33	10.33	22.5
OM2012	1.696	1.332	2.31	0.8	0.72	0.8658	0.9084	0.9473	0.7197	7	5.5	21
OM2011	1.845	1.705	2.698	0.8284	0.8056	0.905	0.9037	0.9165	0.7073	8	6.333	36.6
HARN2014	1.733	1.609	2.137	0.8125	0.8	0.8481	0.9428	1	0.77	8	15	18
HARN2013	2.043	1.889	1.832	0.8642	0.8395	0.7916	0.9644	0.9448	0.6248	18.5	10.33	15
HARN2012	1.609	1.831	1.897	0.8	0.8148	0.8245	1	0.8915	0.7405	15	22	15
HARN2011	1.778	1.479	2.358	0.8047	0.7456	0.8844	0.8456	0.8778	0.8129	10	5	15

OTJ2014	1.523	1.973	1.799	0.7654	0.84	0.755	0.9172	0.899	0.6715	5.333	29	11.5
OTJ2013	1.277	1.55	1.707	0.6939	0.7755	0.7717	0.8965	0.9421	0.6894	4.5	6	9
OTJ2012	1.609	1.523	1.658	0.76	0.7654	0.7704	0.8333	0.9172	0.7499	9	5.333	7.5
OTJ2011	1.735	1.841	1.975	0.8148	0.8281	0.8347	0.9449	0.9002	0.8005	6.75	7.333	9.25

APPENDIX 4. Loadings Matrix values for extracted Principal Components

	PC1	PC 2	PC 3	PC 4
Phylum Firmicutes	LOADINGS			
<i>Streptococcus</i>	0.0060097	0.015581	0.070038	0.034535
<i>Lactococcus</i>	0.013299	0.010273	0.0010616	-0.028358
<i>Brevibacillus</i>	0.02071	0.082062	-0.054274	0.010136
<i>Lysinibacillus</i>	0.029305	-0.020363	0.15551	-0.38771
<i>Paenibacillus</i>	0.0081372	0.036079	-0.013346	0.052516
<i>Bacillus</i>	0.74166	-0.32579	0.089104	0.1734
<i>Unclassified Bacilli</i>	-0.0046433	0.019095	-0.038834	-0.09034
Phylum Actinobacteria				
<i>Arthrobacter</i>	0.0030223	-0.01287	0.02676	-0.029856
<i>Curtobacterium</i>	0.008122	0.0063381	0.011577	-0.006834
<i>Streptomyces</i>	0.00932	0.15111	0.048268	0.010769
<i>Unclassified Actinomycetales</i>	0.0042652	0.040142	-0.040942	-0.054934
Phylum Proteobacteria				
Alphaproteobacteria				
<i>Ochrobactrum</i>	0.012007	-0.0065397	0.035284	0.0036394
<i>Rhizobium</i>	0.024205	-0.019584	0.063551	0.038655

Betaproteobacteria

<i>Achromobacter</i>	0.0038854	-0.012878	0.023707	0.010473
<i>Burkholderia</i>	0.01429	-0.0032358	-0.004771	0.03293

Gammaproteobacteria

<i>Acinetobacter</i>	0.30887	-0.43777	0.22776	0.03672
<i>Azomonas</i>	0.008122	0.0063381	0.011577	-0.006834
<i>Cosenzaea</i>	-0.00059071	0.026231	0.0028788	-0.045744
<i>Enterobacter</i>	0.13287	0.53338	0.74013	0.095357
<i>Escherichia/Shigella</i>	0.029305	0.033826	0.1117	0.05101
<i>Klebsiella</i>	0.083164	0.28661	0.1151	0.11357
<i>Kosakonia</i>	-0.011674	0.070752	0.0021358	0.15392
<i>Pantoea</i>	0.07946	0.15436	-0.2279	0.75846
<i>Stenotrophomonas</i>	0.072226	-0.13499	0.2294	-0.21618
<i>Pseudomonas</i>	0.055454	-0.03682	0.07335	-0.018747
<i>Trabulsiella</i>	0.008122	0.0063381	0.011577	-0.006834
Unclassified <i>Enterobacteriaceae</i>	0.55786	0.50014	-0.46036	-0.3515
Unclassified <i>Pseudomonadaceae</i>	-0.00059071	0.026231	0.0028788	-0.045744
Unclassified <i>Xanthomonadaceae</i>	0.017928	-0.0037881	0.0076442	0.0021248

APPENDIX 5: Number and identity of *T.esculentum* bacterial community identified via the RDP database based on 16S rRNA gene sequences generated using Roche/454 pyrosequencing.

		18T16S	19L16S	19S16S	24L16S	11T16S	11S16S	18L16S	18S16S	11L16S	24S16S
Bacteria	domain	528	64	64	140	1161	107	181	100	226	200
Proteobacteria	phylum	58	22	22	31	338	57	22	5	98	77

[illegible]

Alphaproteobacteria	class	5			1	8					
Caulobacterales	order	3				3					
Caulobacteraceae	family	3				3					
Brevundimonas	genus	3				3					
Rhizobiales	order	2				3					
Brucellaceae	family	2				1					
Ochrobactrum	genus	2				1					
unclassified Rhizobiales						2					
Epsilonproteobacteria	class		2	2					5	2	
Campylobacterales	order		2	2					5	2	
Campylobacteraceae	family		2	2					5	2	
Campylobacter	genus		2	2					5	2	
unclassified Alphaproteobacteria				0	1	2			0	0	
unclassified Gammaproteobacteria						5			0	0	
Firmicutes	phylum	264	2	2	34	620	27		1	1	4
Bacilli	class	264			34	614	26		1		1
Bacillales	order	263			34	608	26		1		
Bacillaceae 1	family	222			19	556	20				
Bacillus	genus	221			19	553	16				
unclassified Bacillaceae 1		1				3	4				
Bacillaceae 2	family					2					
unclassified Bacillaceae 2						2					
unclassified Bacillales		41			15	50	6		1		
Lactobacillales	order										1
Streptococcaceae	family										1
Streptococcus	genus										1
unclassified Bacilli		1				6					
Clostridia	class		2	2					1		3

Cytophagia	class											1	
Cytophagales	order											1	
Cytophagaceae	family											1	
Hymenobacter	genus											1	
Bacteroidia	class												1
Bacteroidales	order												1
Prevotellaceae	family												1
Alloprevotella	genus												1
unclassified Firmicutes							6	1					
unclassified Bacteria		96	25	25	25	150	21	58	26	55	86		
unclassified Root			3	3		262	38	2		22	28		

APPENDIX 6: Loading matrix values for extracted principal Component for pyrosequencing

	PC 1	PC 2
Proteobacteria	0.18997	0.39551
Gammaproteobacteria	0.14655	0.41192
Enterobacteriales	0.16092	0.11537
Enterobacteriaceae	0.16092	0.11537
Enterobacter	0.001044	-0.0033417
Leclercia	0.00062841	0.00069402
Pantoea	-0.0011362	-0.0069665
unclassified_Enterobacteriaceae	0.16038	0.12498
Xanthomonadales	0.0033161	-0.00032836
Xanthomonadaceae	0.0033161	-0.00032836
Stenotrophomonas	0.0017503	0.0023076

unclassified_Xanthomonadaceae	0.0015658	-0.002636
Pasteurellales	-0.021227	0.29438
Pasteurellaceae	-0.021227	0.29438
Haemophilus	-0.020361	0.29135
unclassified_Pasteurellaceae	-0.00086595	0.0030318
Pseudomonadales	0.0004024	-0.00097259
Pseudomonadaceae	0.0004024	-0.00097259
unclassified_Pseudomonadaceae	0.0004024	-0.00097259
Betaproteobacteria	0.029059	-0.0075517
Burkholderiales	0.029158	-0.0089231
Alcaligenaceae	0.029158	-0.0089231
Achromobacter	0.020753	-0.0020797
unclassified_Alcaligenaceae	0.0084052	-0.0068434
Neisseriales	-9.92E-05	0.0013714
Neisseriaceae	-9.92E-05	0.0013714
Neisseria	-9.92E-05	0.0013714
Alphaproteobacteria	0.0057249	-0.0051081
Caulobacterales	0.0023487	-0.003954
Caulobacteraceae	0.0023487	-0.003954
Brevundimonas	0.0023487	-0.003954
Rhizobiales	0.0021942	-0.001942
Brucellaceae	0.00093736	-0.00333
Ochrobactrum	0.00093736	-0.00333
unclassified_Rhizobiales	0.0012568	0.001388
Epsilonproteobacteria	-0.0011428	0.013001
Campylobacterales	-0.0011428	0.013001
Campylobacteraceae	-0.0011428	0.013001
Campylobacter	-0.0011428	0.013001

unclassified_Alphaproteobacteria	0.001182	0.00078787
unclassified_Gammaproteobacteria	0.0031421	0.0034701
Firmicutes	0.42498	-0.11178
Bacilli	0.42212	-0.12198
Bacillales	0.41829	-0.1255
Bacillaceae 1	0.38092	-0.069939
Bacillus	0.37915	-0.07046
unclassified_Bacillaceae 1	0.0017699	0.00052117
Bacillaceae 2	0.0012568	0.001388
unclassified_Bacillaceae 2	0.0012568	0.001388
unclassified_Bacillales	0.036115	-0.056951
Lactobacillales	-9.92E-05	0.0013714
Streptococcaceae	-9.92E-05	0.0013714
Streptococcus	-9.92E-05	0.0013714
unclassified_Bacilli	0.0039249	0.0021521
Clostridia	-0.00084215	0.0059249
Clostridiales	-0.00084215	0.0059249
Clostridiales_Incertae Sedis XI	-0.00074293	0.0045536
Parvimonas	-0.00074293	0.0045536
Lachnospiraceae	-9.92E-05	0.0013714
Moryella	-9.92E-05	0.0013714
unclassified_"Proteobacteria"	0.0097787	-0.016746
Actinobacteria	0.03218	-0.033705
Actinobacteria	0.03218	-0.033705
Actinobacteridae	0.032025	-0.031693
Actinomycetales	0.032025	-0.031693
Streptomyicineae	0.019471	-0.013636
Streptomycetaceae	0.019471	-0.013636

Streptomyces	0.018843	-0.01433
unclassified_Streptomycetaceae	0.00062841	0.00069402
unclassified_Actinomycetales	0.012554	-0.018058
unclassified_Actinobacteria	0.00015448	-0.002012
Fusobacteria	-0.0111	0.16645
Fusobacteriia	-0.0111	0.16645
Fusobacteriales	-0.0111	0.16645
Fusobacteriaceae	-0.011	0.16508
Fusobacterium	-0.011	0.16508
unclassified_ "Fusobacteriales"	-9.92E-05	0.0013714
Bacteroidetes	-0.0027063	0.028594
Flavobacteriia	-0.0025071	0.025111
Flavobacteriales	-0.0025071	0.025111
Flavobacteriaceae	-0.0025071	0.025111
Capnocytophaga	-0.0022072	0.018775
unclassified_Flavobacteriaceae	-0.0002999	0.0063354
Cytophagia	-1.00E-04	0.0021118
Cytophagales	-1.00E-04	0.0021118
Cytophagaceae	-1.00E-04	0.0021118
Hymenobacter	-1.00E-04	0.0021118
Bacteroidia	-9.92E-05	0.0013714
Bacteroidales	-9.92E-05	0.0013714
Prevotellaceae	-9.92E-05	0.0013714
Alloprevotella	-9.92E-05	0.0013714
unclassified_Firmicutes	0.003703	0.0042769
unclassified_Bacteria	0.07712	0.06672
unclassified_Root	0.15622	0.26914

Appendix 7: Percentage Sequence similarity of selected isolates associated with *T. esculentum*

CLADE ONE

JD10_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Enterobacter[58%]
 JD63_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Enterobacter[54%]
 JD28_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Enterobacter[73%]
 JD33_149R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Enterobacter[67%]
 JD30_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Enterobacter[67%]
 JD24_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Enterobacter[66%]
 | JD8_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Citrobacter[59%]
 JD27_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Citrobacter[55%]
 | JD9_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Citrobacter[62%]
 JD49_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Citrobacter[48%]
 JD47_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Enterobacter[87%]
 JD56_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Enterobacter[99%]
 JD14_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%]
 Enterobacter[100%]
 JD16_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Enterobacter[99%]
 JD12_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Enterobacter[99%]
 JD4_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Enterobacter[100%]
 JD46_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Enterobacter[56%]
 JD48_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Enterobacter[92%]
 JD19_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Enterobacter[92%]
 JD18_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Escherichia/Shigella[38%]
 JD13_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Trabulsiella[38%]
 JD26_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Citrobacter[56%]

B13_8F "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%]
 Klebsiella[69%]
 R1_8F "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%]
 Klebsiella[67%]
 B16_8F "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%]
 Klebsiella[64%]
 JD39b_1492 "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Pantoea[100%]
 B7_8F "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%]
 Klebsiella[66%]

CLADE 5

D4_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Xanthomonadales[100%] Xanthomonadaceae[100%] Stenotrophomonas[100%]
 NN1_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Xanthomonadales[100%] Xanthomonadaceae[100%] Stenotrophomonas[100%]
 UZ4_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Xanthomonadales[100%] Xanthomonadaceae[100%] Stenotrophomonas[100%]
 UZ5_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Xanthomonadales[100%] Xanthomonadaceae[100%] Stenotrophomonas[100%]
 D2_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Xanthomonadales[100%] Xanthomonadaceae[100%] Stenotrophomonas[100%]
 D14_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Xanthomonadales[100%] Xanthomonadaceae[100%] Stenotrophomonas[100%]
 D15_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Xanthomonadales[100%] Xanthomonadaceae[100%] Stenotrophomonas[100%]
 NN10_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Xanthomonadales[100%] Xanthomonadaceae[100%] Stenotrophomonas[100%]
 UZ3_1492-R "Proteobacteria"[99%] Gammaproteobacteria[98%] Xanthomonadales[88%] Xanthomonadaceae[84%] Xylella[47%]

CLADE 6

JD11_1492-R "Proteobacteria"[100%] Alphaproteobacteria[100%] Rhizobiales[100%] Rhizobiaceae[97%]
 Rhizobium[97%]
JD61_1492 R "Proteobacteria"[100%] Alphaproteobacteria[100%] Rhizobiales[100%] Rhizobiaceae[100%] Rhizobium[100%]

23_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[100%] Planococcaceae[100%] Lysinibacillus[100%]

D10_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[100%] Bacillus[100%]

XY4_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[99%] Bacillus[98%]
 UZ8_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[100%] Bacillus[100%]
 UZ7_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[100%]
 Bacillus[100%]
 D5_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[100%] Bacillus[100%]
 6_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[99%] Bacillus[99%]

NN8_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[100%] Bacillus[100%]
 NN9_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[100%] Bacillus[100%]
 15_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[100%] Bacillus[100%]
 25_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[89%] Bacillus[72%]
 b1.scf Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[93%]
 Bacillus[93%]
 R5_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[100%] Bacillus[100%]
 D6_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[100%] Bacillus[100%]
 UZ6_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[92%] Bacillus[87%]
 XY3_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[100%] Bacillus[100%]
 UZ12_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[100%]
 Bacillus[100%]
 UZ13_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[100%]
 Bacillus[100%]
 UZ10_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[100%]
 Bacillus[100%]

CLADE 7

D7_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Enterobacter[68%]
 NN6_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Klebsiella[75%]
 NN4_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] "Enterobacteriales"[100%] Enterobacteriaceae[100%] Enterobacter[95%]

CLADE 8

B8_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Pseudomonadales[100%] Moraxellaceae[100%] Acinetobacter[100%]

B15_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Pseudomonadales[100%] Moraxellaceae[100%] Acinetobacter[100%]

CLADE

9

D16_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Pseudomonadales[100%] Moraxellaceae[100%] Acinetobacter[100%]

D21_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Pseudomonadales[100%] Moraxellaceae[100%] Acinetobacter[100%]

D1_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Pseudomonadales[100%] Moraxellaceae[100%] Acinetobacter[100%]

NN13_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Pseudomonadales[100%] Moraxellaceae[100%] Acinetobacter[100%]

UZ2_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Pseudomonadales[100%] Moraxellaceae[100%] Acinetobacter[100%]

UZ9_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Pseudomonadales[100%] Moraxellaceae[100%] Acinetobacter[100%]

D8_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Pseudomonadales[100%] Moraxellaceae[100%] Acinetobacter[100%]

D20_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Pseudomonadales[100%] Moraxellaceae[100%] Acinetobacter[100%]

UZ15_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Pseudomonadales[100%] Moraxellaceae[100%] Acinetobacter[100%]

UZ11_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Pseudomonadales[100%] Moraxellaceae[100%] Acinetobacter[100%]

D9_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Pseudomonadales[100%] Moraxellaceae[100%] Acinetobacter[100%]

D11_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Pseudomonadales[100%] Moraxellaceae[100%] Acinetobacter[100%]

D13_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Pseudomonadales[100%] Moraxellaceae[100%] Acinetobacter[100%]

D17_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Pseudomonadales[100%] Moraxellaceae[100%] Acinetobacter[100%]

NN7_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Pseudomonadales[100%] Moraxellaceae[100%] Acinetobacter[100%]

CLAD 2

B6_8F "Proteobacteria"[100%] Gammaproteobacteria[100%] Pseudomonadales[100%] Moraxellaceae[100%] Acinetobacter[100%]

B17_8F "Proteobacteria"[100%] Gammaproteobacteria[100%] Pseudomonadales[100%] Moraxellaceae[100%]
Acinetobacter[100%]

B18_8F "Proteobacteria"[100%] Gammaproteobacteria[100%] Pseudomonadales[100%] Moraxellaceae[100%] Acinetobacter[100%]

B19_8F "Proteobacteria"[100%] Gammaproteobacteria[100%] Pseudomonadales[100%] Moraxellaceae[100%]
Acinetobacter[100%]

CLADE3

R8_8F "Proteobacteria"[100%] Betaproteobacteria[100%] Burkholderiales[100%] Burkholderiaceae[100%] Burkholderia[100%]

JD58_1492-R "Proteobacteria"[100%] Gammaproteobacteria[100%] Pseudomonadales[100%] Pseudomonadaceae[100%] Pseudomonas[98%]

JD34_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[100%]

Bacillus[100%]

JD61_1492 R "Proteobacteria"[100%] Alphaproteobacteria[100%] Rhizobiales[100%] Rhizobiaceae[100%] Rhizobium[100%]

D6_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[100%] Bacillus[100%]

B10_8F Actinobacteria[100%] Actinobacteridae[100%] Actinomycetales[100%] Micrococcineae[100%] Micrococcaceae[100%] Arthrobacter[100%]

CLADE 4

JD15_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[100%]

Bacillus[100%]

JD23_1492-R_B04C.scf Root[100%] Bacteria[100%] Firmicutes[100%] Bacilli[100%] Bacillales[100%] Planococcaceae[100%] Lysinibacillus[100%]

JD25_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[99%] Bacillaceae 1[80%] Bacillus[67%]

JD37_1492-R Firmicutes[100%] Bacilli[100%] Bacillales[100%] Paenibacillaceae 1[100%] Paenibacillus[91%]

R6_8F Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[100%] Bacillus[100%]

B12_8F Firmicutes[100%] Bacilli[100%] Bacillales[100%] Bacillaceae 1[100%] Bacillus[100%]