Seedborne bacteria of onion: a study on pathogenicity and diversity

Submitted by

Mamogwasha Vallry Moloto

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

In the Faculty of Natural and Agricultural Sciences, Department of Plant and Soil Science, University of Pretoria, Pretoria, South Africa

Supervisor: Prof. T. Coutinho

Co-supervisors: Dr. T. Goszczynska

Prof. L. du Toit

February 2022

Declaration

I, Mamogwasha Vallry Moloto, declare that the thesis I hereby submit for the Philosophiae Doctor in Plant Pathology at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

SIGNATURE ………………….

DATE:………………………..

Acknowledgements

- To my mother, Linah, my brother Abram and my daughter Lehlogonolo thank you for the support and encouragement that i received from you guys. When I wilted under pressure of work you kept me going in times of misery.

- My supervisor. Prof. Teresa Coutinho at the University of Pretoria, Dr Teresa Goszczynska at Agricultural Research Council: Plant Protection Research of South Africa and Prof. Lindsey du Toit at Washington State University for all their input into my thesis. I particularly valued their unwavering and insightful guidance of my work that made me feel at home and removed any doubt that I would attain the target.

- I now appreciate my limitations with humbleness and peace and accept that when stuck always seek help from people who can provide answers, My heartfelt thanks also go to the following people, Margot Otto, Rian Pierneef and Solize Vosloo.

- To my colleagues at Agricultural Research Council: Plant Health Protection at Bacterial Disease Unit for their support, especially Mr Harry Boroko for technical assistance.

- To my friends Adoration Shubane, Jerida Mamabolo and Refilwe Ngoepe for their support and making sure that I do not crack through the journey of my research and when writing up the thesis.

- To the late Mapula Motana, I know this thesis comes just too late, but I know you would have been proud of me. These thesis is dedicated to you, I wish you were here to witness all of this, unfortunately you gone too soon. May your soul Rest in Peace.

- Plant Phytobacteriology Group (PPG) at the Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, South Africa.

- The Agricultural Research Council: Plant Protection Research of South Africa and University of Pretoria for financial support.

Preface

Onions (*Allium cepa* L.) are cultivated worldwide and are an important component of countless cuisines, lending varied flavours to meals. In South Africa, onions are grown by both commercial and small-scale farmers and are considered the third most important vegetable crop after potatoes and tomatoes. South Africa is one of the biggest producers of onion seed in the world. In the 2014/2015 growing season 724 tons of onion seed were produced, of which 657 tons were for the export market.

Pantoea ananatis, *Pantoea. allii*, *Pantoea agglomerans*, *Pseudomonas syringae* pv. *porri*, *Pseudomonas syringae* pv. *allii* and *Xanthomonas axonopodis* pv. *allii* are seedborne bacterial pathogens of onions. Bacterial blights of onion can be caused by *Pseudomonas syringae* pv. *porri*, *Pseudomonas syringae* pv. *allii* and *Xanthomonas axonopodis* pv. *allii* and symptoms caused by the different pathogens are indistinguishable in the field. Bacterial pathogens reduce seed quality and yield to onion producers. The use of pathogen-free seed is recommended to prevent outbreaks of these diseases. Significant progress has been made in developing reliable, sensitive and specific techniques for pathogen detection, but epidemics resulting from seedborne inocula continue to occur.

The first chapter presents a literature review on bacteria associated with onion seeds, their detection, identification and pathogenicity determinants. The review also focuses on the various techniques used to describe bacterial communities. A discussion on the application and significance of whole genome sequencing and analysis is included.

Bacterial pathogens of onion plants presence in seed cause substantial losses to onion producers. **The second chapter** presents the characterisation of *Pseudomonas syringae* strains isolated from onion plants and seeds. Biolog GN III and multilocus sequence typing (MLST) based on partial sequences of four housekeeping genes (*cts*, *gapA*, *gyrB* and *rpoD)* were used to differentiate *P. syringae* pathovars. Pathogenicity tests were performed on onion (cv. Granex 33), chive (*Allium schoenoprasum* cv. Grasiue), leek (*Allium porrum* cv. Giant Italian) and spring onion (*Allium fistulosum* cv. Salotte) plants to assess differences in host range among isolates.

Seeds can carry diverse microbial communities, which may have beneficial or harmful effects on plant growth and health. **The third chapter** describes an analysis of the bacterial

communities associated with onion seed lots of a single cultivar using culture-dependent and culture-independent methods.

In the fourth chapter, the full genome sequences of *P. agglomerans* strains BD 1212 (non-pathogenic on onion) and BD 1274 (pathogenic on onion), isolated from onion seeds, were sequenced, assembled and annotated. The full genomes were then used for comparative genomic analysis. Genes that differentiated the non-pathogenic from pathogenic strains were discussed in order to answer questions related to the adaptation mechanisms and fitness of the non-pathogenic strain and potential virulence mechanisms of the pathogenic strain.

In this study, the research described in each chapter was done and described independently to avoid redundancy among chapters. It is my hope that the results of these studies on seedborne pathogens of onion will contribute to a better understanding of their emergence, survival, colonisation and virulence on onion.

Table of Contents

Declaration 2
Acknowledgements 3
Preface 4

Chapter 1
Bacteria associated with onion seed: a review of seedborne bacterial pathogens, factors related to their pathogenicity and bacterial communities in seed

1.1 Introduction 10
1.2 Seedborne bacterial pathogens of onion 11
1.2.1 Bacterial leaf blight, seed stalk necrosis and centre rot of onion 12
1.2.1.1 Symptoms 12
1.2.1.2 Identification and detection 13
1.2.3 Bacterial leaf blight of onion 13
1.2.3.1 Symptoms 15
1.2.3.2 Detection and identification of *Xanthomonas axonopodis* pv. *allii* 15
1.2.3.3 Detection and identification of *Pseudomonas syringae* pv. *porri* 16
1.2.3.4 Disease cycle of seedborne pathogens of onion 16
1.3 Current management strategies of seedborne pathogens 17
1.4 The mechanisms for tolerance in rhizosphere bacteria that provides a synthesis of the types of potential treatment used for bacterial pathogens 19
1.5 Methods used to describe microbial communities in and on seed 21
1.5.1 Bacterial communities in and on seed 22
1.5.1.1 Culture-dependent methods 22
1.5.1.2 Culture-independent methods 22
1.6 Pathogenicity and virulence factors of seedborne plant pathogenic bacteria 24
1.6.1 Secretion systems 25
1.6.2 Phytotoxins produced by *Pseudomonas* species 29
1.6.3 Quorum sensing and biofilm production 30
1.6.4 Pili, fimbriae, and flagella 30
1.7 Additional factors associated with pathogenicity 31
1.8 Bacterial avirulence genes 32
1.9 Whole genome sequencing 33
1.10 Conclusion 37
1.11 References 38

1.12 Figures 57

Chapter 2
A new pathovar of *Pseudomonas syringae*, pathovar *allii*, isolated from onion plants exhibiting symptoms of blight

2.1 Abstract 59
2.2 Introduction 59
2.3 Materials and Methods 63
2.4 Results 64
2.5 Discussion 65
2.6 References 68
2.7 Tables 72
2.8 Figures 79
2.9 Online supplementary materials 82

Chapter 3
Community analysis of bacteria associated with onion (*Allium cepa* L.) seed lots produced in South Africa

3.1 Abstract 87
3.2 Introduction 87
3.3 Materials and methods 89
3.4 Results 94
3.5 Discussion 96
3.6 References 99
3.7 Tables 102
3.8 Figures 112
3.9 Supplementary information 117

Chapter 4
Comparative genomics of pathogenic and non-pathogenic *Pantoea agglomerans* strains isolated from onion seeds

4.1 Abstract 122
4.2 Introduction 122
4.3 Material and methods 124
4.4 Results 127
4.5 Discussion 132
4.6 Conclusion 137

4.7 References 138
4.8 Tables 147
4.9 Figures 154
4.10 Supplementary information 158

General discussion 167

Summary 170

Appendix A: Seed stand for onion for onion production (Plant Imptovement Act (1976) edited in 2018) 173

Chapter 1

Bacteria associated with onion seed: a review of seedborne bacterial pathogens, factors related to their pathogenicity and bacterial communities in seed

1.1 Introduction

Onion (*Allium cepa* L.) is a member of the family Alliaceae, belonging to the genus *Allium* which is believed to originate in Asia (Brewster, 2008; Fabrice et al., 2018; Pareek et al., 2017). Onion is reported to contribute to human health and is useful for flavouring food and as a herb (Brewster, 2008; Fabrice et al., 2018; Pareek et al., 2017). Study has shown that onion and leek (*A. ampeloprasum* L. var. *porrum* L.) are the most common edible *Allium* species grown worldwide (Brewster, 2008; Pareek et al., 2017). In South Africa, onions are considered the third most important crop (Department of Agriculture, Forestry and Fisheries, 2010; The National Agricultural Directory, 2011). South Africa is one of the biggest producer of onion seeds worldwide. Total volumes of 724.80 tons of onions were produced in the 2014/2015 growing season, of which 657.07 tons were for the export market (SANSOR Annual Report, 2015).

The vegetable seed market in South Africa has shown a growth of 7% in 2019 compared to 2018 (SANSOR Annual Report, 2021). Research showed that sales of the top seeds continued to increase, with Short Day Onion seed (Fig. 1.1) leading the sales board in South Africa (www.sansor.org). The impact of climate change on seed production is a huge concern and it is clear that major events like droughts, floods and especially heat waves are causing havoc in seed production around the globe. In 2019, onion, carrot, bunching onion seeds production (Fig. 1.2) were on the increase in volume and value as compared to 2018 (SANSOR Annual Report, 2021). South African National Seed Organization (SANSOR) reported that there is growing concern in the seed production industry about the ever-increasing quality standards required by the seed market (www.sansor.org). This includes better germination, better variety purity, better uniformity, etc. (www.sansor.org).

Onions are grown commercially in South Africa, mainly in the Free State, Western Cape, Limpopo, North West and Northern Cape Provinces (Department of Agriculture, Forestry and Fisheries, 2015). The Klein Karoo region in the southern Cape, with its dry, warm climate, is favourable for the production of onion seed for clients worldwide (SANSOR Annual Report, 2021). Unfortunately, most parts of the Northern Cape, Western Cape, Eastern Cape, Karoo and Klein Karoo are still being affected by the enduring drought (SANSOR Annual Report, 2021). The onion industry uses fresh produce markets, informal markets, processors, and direct selling to wholesalers and retailers as marketing channels.

Onions are produced mainly for their use as food and for oil extraction (Department of Agriculture, Forestry and Fisheries, 2010; Fabrice et al., 2018; Pareek et al., 2017). Randle et al. (1998) emphasised the importance of choosing the appropriate cultivar when planting onions for food as cultivars differ in their taste. Onions need to be grown under cool conditions, with an optimum temperature between 22 to 28°C for good vegetative growth (Starke Ayres, 2015; www.starkeayres.co.za). Onion seed production requires low-humidity ambient conditions during the spring and summer (www.starkeayres.co.za)

Seeds are carriers of plant pathogens worldwide and may even be accountable for the introduction of quarantine pathogens and the re-emergence of old diseases (Dutta et al., 2014; Gitaitis & Walcott, 2007). The risk of introducing unexpected seedborne pathogens also increases as global trade grows (Gitaitis & Walcott, 2007; Schaad et al., 2014). Infested seeds may be responsible for disease outbreaks resulting in serious economic losses. Gitaitis & Walcott (2007) proposed that research institutes and governments should work together to provide pathogen free-seeds. Seed can be assessed only if seed health assays have been developed for the specific pathogen(s) of concern. Infested seed lots should be either destroyed or treated with fungicides or physical methods to kill pathogens (Schaad et al., 2014). Significant progress has been made in developing reliable, sensitive and specific techniques for detection of seedborne bacterial pathogens, but technical challenges remain. For example, seeds can be contaminated with both pathogenic and non-pathogenic strains of the same bacterial species. Methods to differentiate the target pathogen(s) from non-target bacteria in seed health assays are required in such situations (Dutta et al., 2014).

This review focuses on important seedborne pathogens of onion, with a focus on *Pseudomonas* and *Pantoea* species, their distribution and ability to cause disease symptoms on the host. Microbial diversity in seeds, virulence and avirulent factors and the role played by horizontal gene transfer in the acquisition of these factors in plant pathogenic bacteria are also discussed.

1.2 Seedborne bacterial diseases of onion

Seedborne pathogens are economically important because they can act as sources of inocula for many important vegetable crops (Dutta et al., 2014). Pathogens of onion are problematic for onion growers around the world because they can cause foliar diseases and bulb rots (Gitaitis & Walcott, 2007; Dutta et al., 2014; Zaid et al., 2012). Below, we describe important seedborne pathogens of onion plants and seeds. This includes the causal agents of bacterial centre rot of

onion, *Pantoea ananatis* and *Pantoea allii* (Brady et al., 2011; Goszczynska et al., 2006; Walcott et al., 2002); leaf and seed stalk necrosis of onion, *P. agglomerans* (Goszczynska et al., 2006); and bacterial blight of leek and onion, *Pseudomonas syringae* pv. *porri* (Myung et al., 2011; Nobel et al., 2006, Samson et al., 1998) and *Xanthomonas axonopodis* pv. *allii* (Kodata et al., 2000; Roumagnac et al., 2004).

1.2.1 Bacterial leaf blight, seed stalk necrosis and centre rot of onion

Bacterial diseases of onion include centre rot caused by *Pantoea ananatis* and *Pantoea allii*; and leaf blight and seed stalk necrosis caused by *Pantoea agglomerans* (Brady et al., 2011; Edens et al., 2006; Gitaitis & Gay, 1997; Hattingh & Walters, 1981). Gitaitis & Gay (1997) reported bacterial centre rot, induced by *P. ananatis,* for the first time in the USA (Gitaitis & Gay, 1997). Afterward it was described in onion production field areas of Colorado, Michigan and New York (Carr et al., 2010; Gitaitis et al., 2002; Schwartz & Otto, 1998, 2000). Since the seed related with the first outbreak of centre rot in Georgia was produced in South Africa, it was concluded that infested onion seed (Walcott et al., 2002) introduced centre rot. However, symptoms of centre rot of onion had never been witnessed in South Africa even though *P. ananatis* was identified from onion seed lots obtained from small-scale farmers (Brady et al., 2011; Goszczynska et al., 2006). In 1981, Hattingh and Walters reported leaf and seed stalk blight in South Africa, which induced symptoms similar to that of centre rot. The authors identified the pathogen as *P. agglomerans* (= *Erwinia herbicola*). In 2012, a disease identical to that caused by *P. agglomerans* in South Africa was reported in Korea (Kim & Choi, 2012). Leaf blight and bulb decay caused by *P. allii* was first reported in 2011 on onion plants and seed lots in USA and South Africa, respectively (Brady et al., 2011; Goszczynska et al., 2006).

1.2.1.1 Symptoms

Centre rot of onion typically affects the centre leaves of the plant (Gitaitis & Gay, 1997). The leaves become water-soaked, tan or brown. As the disease progresses, all leaves may wilt. Symptoms often spread into a bulb, which can become soft and may produce a foul odour. Infected seed stalks exhibit symptoms similar to that occurring in leaves (Brady et al., 2011; Goszczynska et al., 2006; Hattingh & Walters, 1981; Gitaitis & Gay 1997).

1.2.1.2 Identification and detection

Pantoea ananatis, *P. allii* and *P. agglomerans* are Gram negative, rod-shaped bacteria that produce a yellow pigmentation when grown on tryptone glucose extract agar (TGA) (Difco). They are catalase positive, oxidase negative and facultative anaerobes. *Pantoea agglomerans* differs from *P. ananatis* and *P. allii* by not producing indole. Furthermore, the ability of *P. allii* to use adonitol is a useful phenotypic feature to distinguish it from *P. ananatis* (Brady et al., 2011).

Previously, identification of *Pantoea* species was based on biochemical characteristics and commercialised phenotypic identification systems (Gavini et al., 1989). Today, identification of *Pantoea* strains is based on PCR analyses with species-specific primers (Gitaitis et al., 2002; Walcott et al., 2002); sequencing of the 16S rRNA gene (Cother et al., 2004; Coutinho et al., 2002; Medrano & Bell, 2007; Schmid et al., 2003), amplified fragment length polymorphism (AFLP) analysis (Brady et al., 2007; Bulletin, European and Mediterranean Plant Protection Organization (EPPO), 2016; Goszczynska et al., 2007; Roumagnac et al., 2004) and multilocus sequence analysis (MLSA) (Brady et al., 2008, 2011). In 2008, Brady et al. developed a MLSA scheme based on sequencing of four housekeeping genes, *atpD*, *gyrB*, *infB* and *rpoB* for characterising *Pantoea* species. Their results distinguished all the species from each other and these results were further supported by DNA-DNA hybridization. The MLSA approach was used to describe new *Pantoea* species and reclassify other species that fell in the core *Pantoea* cluster (Brady et al., 2008, 2011).

1.2.3 Bacterial leaf blight of onion

Roumagnac et al. (2004) and Moloto et al. (2017) reported that *X. axonopodis* pv. *allii* and *P. syringae* pv. *porri* cause leaf blight of onion. Bacterial blight of onion, also known as Xanthomonas leaf blight of onion (Kodata et al., 2000), was first described in Hawaii in 1978 (Alvarez et al., 1978). Later, the disease was reported in onion production field in the East Caribbean, South America, Asia, Brazil, South Africa, United States and Rèunion Island (Kadota et al., 2000; Nunez et al., 2000; Pereira & Tebaldi, 2013; Picard et al., 2008; Roumagnac et al., 2000; Sander et al., 2003; Serfontein, 2001). *Xanthomonas axonopodis* pv. *allii* infects chives, garlic, leek, onion, shallot and welsh onion (Bulletin EPPO, 2016; Kadota et al., 2000; Picard et al., 2008; Roumagnac et al., 2004).

The formal way of classifying pseudomonad plants pathogens began with the name *Pseudomonas syringae* because the original type strain was a pathogen of lilacs within the genus *Syringa* (Young 2010). Strains were formerly distinguished as *P. syringae* complex if they are fluorescent pigment on King's B medium, produce a hypersensitive reaction on tobacco; produce levan on sucrose medium, utilise arginine, negative for oxidase production activity and cause potato rot (LOPAT tests) (Lelliott et al., 1966; Moloto et al., 2017). *Pseudomonas syringae* has a wide host range with more than 180 plant species comprising flowers, fruit trees, and different crops (Kaluzna et al., 2012). Research have shown that *P. syringae* is the most significant and economically important plant pathogenic bacterial species (Mansfield et al., 2012).

Pseudomonas syringae species are re-classified into nine various genomospecies based on DNA-DNA-hybridization and 13 phylogenetic groups based on multilocus sequence typing (MLST) (Barret et al., 2015; Bull et al., 2011; Gardan et al., 1999; Moloto et al., 2017). Genomospecies can be given an official name only when the phenotypic characteristics that separate the strain from other can be described (Wayne et al., 1987). *Pseudomonas cannabina* and *P. tremae* were elevated to species level because their unique carbon source utilisation (phenotypic features) were reported that enabled these species to be differentiated from other species of *P. syringae* (Bull & Koike, 2015; Marcelletti & Scortichini, 2014; Moloto et al., 2017). Berge et al. (2014) described 13 phylogroups that belong to *P. syringae sensu lato*. Phylogroups 1, 2, 3, 5, 6 and 10 consist of the strains that belong to *P. syringae* complex and phylogroups 4, 7, 8, 9, 11, 12 and 13 contains many strains that have been allocated to species such as *P. viridiflava* and *P. cichorii* (Baltrus et al., 2014; Berge et al., 2014; Morris et al., 2008).

Goto (1972) first recorded *P. syringae* as a disease of *Allium* in Japan. The author described leaf spot symptoms of onions as lesions with a beige and a water-soaked margin. In 2011, Myung et al. described a bacterial leaf spot disease of onion induced by *P. syringae* pv. *porri* for the first time in Korea. In 2012, the same pathogen was reported in Georgia, USA, and then in Queensland, Australia in 2014 (www.planthealthaustralia.com.au). Symptoms of bacterial blight of onion caused by *X. axonopodis* pv. *allii* are identical to those of leaf blight of leek caused by *P. syringae* pv. *porri* (Moloto et al., 2017; Noble et al., 2006; Samson et al., 1998). Lately, Moloto et al. (2017) reported a new pathovar of *Pseudomonas*, *P. syringae* pv. *allii*, showing blight symptoms on onion in South Africa. Recently, Tsuji & Takakawa (2018) reported a new pathovar of *Pseudomonas*, *P. syringae* pv. *alliifistulosi* pv. nov., causing

bacterial leaf spot of onions. The authors distinguished *P. syringae* pv. *alliifistulosi* pv. nov. from *P. syringae* pv. *porri* by rep-PCR. Additionally, MLSA analysis on housekeeping genes and *hrp* genes which encode the type-III secretion system indicated that the strains of *P. syringae* pv. *alliifistulosi* pv. nov. group individually from *P. syringae* pv. *porri* (Tsuji & Takakawa, 2018).

1.2.3.1 Symptoms

Symptoms of onion leaf blight consist of water-soaked lesions, which become decolorise and then die (Serfontein, 2001; Roumagnac et al., 2004). Lesions increase into long, chlorotic strips and water-soaked spot become more obvious on the flattened side of older leaves. Blighting of leaves and tip death reduce the plant photosynthetic area and decrease the bulb size (Serfontein, 2001; Roumagnac et al., 2004). Symptoms of leaf blight in other *Allium* species are undistinguishable to those on onion plants (Bulletin EPPO, 2016; Picard et al., 2008; Roumagnac et al., 2004). Roumagnac et al. (2004) reported that, leaf dieback can occur when the disease is critical, resulting in reduction of bulb size. Severe reduction in onion bulb size has been reported from South Africa (Bulletin EPPO, 2016; Serfontein, 2001). As a results, profit losses vary between 10% to 50% have been reported in the USA (Nunez et al., 2002; Sander et al., 2003; Schwartz & Otto, 2000).

1.2.3.2 Detection and identification of *Xanthomonas axonopodis* pv. *allii*

Detection of *X. axonopodis* pv. *allii* is performed by plating a bacterial suspension obtained from diseased plant tissue extracts on selective media, e.g., Milk-Tween (MT) medium (Goszczynska & Serfontein, 1998). Identification from the genus to the species level is done by sequencing the gyrase B (*gryB*) gene (Parkinson et al., 2007). Amplified fragment length polymorphism (AFLP) and repetitive sequence-based PCR (rep-PCR) (Gent et al., 2004; Humeau et al., 2006; Roumagnac et al., 2004) can be used for preliminary screening of the *Xanthomonas* isolates. A real time PCR protocol designed by Robène et al. (2015) can be used directly on ground onion seeds for preliminary evaluation of *X. axonopodis* pv. *allii*. However, identification of the pathogen is not complete without performing a pathogenicity test to confirm that the bacterial strains are able to induce symptoms on the host plants. If the pathogen can be re-isolated from inoculated plants, Koch's postulates are fulfilled.

1.2.3.3 Detection and identification of *P. syringae* pv. *porri*

Detection of *P. syringae* pv. *porri* is performed by plating a bacterial suspension obtained from diseased plant tissue extracts on selective media, e.g., King's B (KB) medium (King et al., 1954). The LOPAT tests (Lelliott et al., 1966) and carbon source utilisation test (Young & Triggs, 1994) are used to distinguish pathovars of *P. syringae*. Rep-PCR DNA fingerprinting can also be used to separate *P. syringae* pathovars (Koike et al., 1999; Louws et al., 1994; Noble et al., 2006; Tsuji & Takikawa, 2018).

Multi-locus sequence typing is done by the sequence analyses of the following housekeeping genes: glyceraldehyde-3-phosphate dehydrogenase (*gapA*), gyrase B (*gyrB*), citrate synthase (*cts*) and sigma factor 70 (*rpoD*) (Berge et al., 2014; Hwang et al., 2005; Moloto et al., 2017; Myung et al., 2011). Berge et al. (2014) published a guide to study the diversity and method to characterise *P. syringae* isolates. They proposed using *cts* sequences as a quick and accurate means of characterising new strains to the pathovar level. However, achievement of Koch's postulates is still required to identify the pathovar of bacterial strains (Koike et al., 1999; Noble et al., 2006, Tsuji & Takikawa, 2018).

1.2.3.4 Disease cycle of seedborne bacteria of onion

The interaction between the pathogen/s, the host plant and the environment results in the occurrence of a disease outbreak (Agrios, 2005). The first requirement for the initiation of disease is for the pathogen to come into contact with the host (Agrios, 2005). Planting non-certified seeds (which could be contaminated) or infected transplants can lead to bacterial disease outbreaks. Figure 1.1 discusses the cycle of a typical bacterial seedborne disease of onion. Briefly, the pathogen survives in the seeds and, potentially, in plant debris that has not been destroyed (Black et al., 2012; Gitaitis & Walcott, 2007; Schwartz & Gent, 2007). When infected seeds are planted in the field or glasshouse, bacteria carried on or within the infected seeds can contaminate the surface of the growing cotyledon when each seed germinates (Agarwal et al., 2019).

The seedling infected with the pathogen can spread the bacterium through irrigation water and rains (Black et al., 2012; Gitaitis & Walcott, 2007; Schwartz & Gent, 2007). The pathogen enters wounded plants via foliage or scape (seed stalk) through stomata or through wounds from hail damage, feeding insects, sunscald, freezing, etc. (Agarwal et al., 2019; Gitaitis & Walcott, 2007; Schwartz & Gent, 2007). Infected plants produce water-soaked lesions disease is favoured by rain, warm to hot weather (depending on the genus and species of bacterium

infecting onion plants), and develops rapidly at temperatures between 24 and 30°C (Agarwal et al., 2019; Gitaitis & Walcott, 2007; Schwartz & Gent, 2007). Infection can spread on or between plants by splashing water, insects (e.g., thrips), or mechanically from plants rubbing against each other, including dispersal from infected leaves or scapes (seed stalks) onto the flowers and developing seed in the umbels (Agarwal et al., 2019). The pathogen can multiply on the floral surface and can be moved from flower to flower by splashing water or feeding insects (Agarwal et al., 2019; Gitaitis & Walcott, 2007). After harvest, the bacteria survive commercial processing, harvesting, cleaning and storage of the seed, and then persist in the seeds (Agarwal et al., 2019).

1.3 Current management strategies of seedborne pathogens

Management of plant diseases is essential for most crops, and is critical for the production of high quality seed (du Toit, 2004; Mancini & Romanazzi, 2014). Plant pathogens might reduce the quantity and quality of the seeds harvested, and some can be seedborne pathogens of the host plants. For onions such as *Xanthomonas axonopodis* pv. *allii*, *Pantoe*a spp. and *Pseudomonas porri* (Goszczynska et al., 2006; Moloto et al., 2017; Myung et al., 2011; Roumagnac et al., 2004). In today's era of globalisation and free trade, seed accounts for disseminating some plant pathogens across huge distances and political borders (du Toit, 2004; Danesh et al., 2014; Gitaitis & Walcott, 2007).

Seed health testing has an important influence on disease management (Gupta & Kumar, 2020). In general, methods used for managing diseases of seeds crops include exclusion of pathogens from regions of seed production, eradication of pathogens from seed crops, protection of seed crops with fungicides or bactericides, reduction in disease pressure using cultural practices, and physical treatment (Danesh et al., 2014; du Toit, 2004; Gupta & Kumar, 2020). Seed treatments sometimes offer an effective means of eradicating or reducing the prevalence of seedborne pathogens, which is especially important when seeds are planted for seed production (du Toit, 2004; Gupta & Kumar, 2020; Mancini & Romanazzi, 2014).

Eradication methods are applied directly against the pathogen to the host plants (Danesh et al., 2014; Gupta & Kumar, 2020). Practical eradication procedures include fumigation of seed storage houses, heat treatment, solarisation and burning or removal of infected plant residues. The disinfection of infected seeds by physical seed treatments (e.g., by heating with hot water or steam) can be an effective measure to eliminate primary inoculum of a pathogen and to

prevent disease establishment in fields (Danesh et al., 2014; Mancini & Romanazzi, 2014; Gupta & Kumar; 2020). Chemical seed treatments (e.g., with chlorine) can eradicate or reduce inoculum of some pathogens on seeds (Danesh et al., 2014; du Toit, 2004; Gupta & Kumar, 2020).

Cultural practices can contribute highly to the control of many plant diseases (Danesh et al., 2014; Gupta & Kumar, 2020). The aim of cultural disease management practices is to restrict growth of the pathogens and offer favourable environmental conditions for crop development, which results in good plant health, and less favourable conditions for the pathogens (Danesh et al., 2014; Gupta & Kumar, 2020; Gitaitis & Walcott, 2007). Crop inspections and roguing of symptomatic plants or symptomatic alternative hosts of plant pathogens can reduce disease pressure in a seed crop (du Toit, 2004; Gupta & Kurmar, 2020; Gitaitis & Walcott, 2007). A minimum 3-year crop rotation with non-host crops is advisable to greatly reduce populations of the pathogens (Danesh et al., 2014; Gupta & Kurmar, 2020). Crop rotation is particularly effective for management of foliar pathogens, but longer rotations are needed for soilborne pathogens.

Current management practices for bacterial blight of onion involve the use of pathogen-free seeds, removal of infected crop residues, crop rotation and applications of copper-based bactericides (Black et al., 2012; www.gardeningknowhow.com). Seeds should be produced in areas where the pathogens of major concern are less able to establish or achieve critical population size during seed development (Agarwal et al., 2019; Danesh et al., 2014; Gitaitis & Walcott, 2007). Use of certified onion seed lots that have been tested for specific seedborne pathogens is encouraged to avoid introduction of inoculum of these seedborne pathogens, such as *Pantoea* spp., into production fields, since onion cultivars with resistance to *Pantoea* species are not available (Agarwal et al., 2019; Gitaitis & Walcott, 2007).

Overhead irrigation should be avoided as it promotes bacterial spread by splashing compared with sub-surface or drip-irrigation (www.gardeningknowhow.com; Agarwal et al., 2019; Gitaitis & Walcott, 2007). The implementation of successful weed management strategies is important in reducing *P. ananatis* inoculum in fields as some weeds are alternative hosts to this bacterial pathogen and weeds also increase the density of the canopy of a crop, which reduces air flow, increasing humidity and creating more favourable conditions for infection of onion plants by bacteria (Agarwal et al., 2019; Danesh et al., 2014; Gitaitis & Walcott, 2007). Controlling thrips can be an effective management strategy to reduce centre rot incidence as

these insects can serve as vectors of *Pantoea* spp. and play an important role in bacterial transmission to onion plants (Agarwal et al., 2019; Dutta et al., 2014). Management of centre rot in onion fields also tends to be based on applications of copper bactericides, often mixed with an ethylene bis-dithiocarbamate fungicide (EBDC), such as mancozeb, which growers may apply weekly as a protectant against bacterial infection (Agarwal et al., 2019).

In the past three to four decades, many studies have focused on controlling plant pathogens by application of biological control agents (e.g., Compant et al., 2005; Kim et al., 2012; Johnson & Stockwell, 1998). Applications of biocontrol microorganisms were reported to have the potential to improve seedling health by protecting seedlings against seedborne plant pathogens (Berg, 2009). However, Barret et al. (2015) stated that potential varying efficacy of these biological control treatments usually is not be acceptable to farmers. Consequently, the search for biocontrol agents that provide adequate and consistent protection against plant diseases across agricultural production continues (Barret et al., 2015). Links et al. (2014) reported that a *P. agglomerans* strain isolated in their study had the capacity of being a biocontrol agent when applied to *Triticum* and *Brassica* seeds by protecting the seeds from microorganisms associated with spoilage, such as *Alternaria* spp. Furthermore, Bulgarelli et al. (2013) stated that bacterial communities on flowers and seeds might serve as reservoirs for biological control of microbial pathogens of crops, e.g., pumpkin (*Cucurbita*) pathogens. To date, no biological control treatments have proven to be highly effective against seedborne bacterial pathogens of onion.

1.4 The mechanisms for tolerance in rhizosphere bacteria that provides a synthesis of the types of potential treatment used for bacterial pathogens

Plant growth promoting rhizobacteria (PGPR) are a group of bacteria found in the rhizosphere (Ahmad et al., 2008; Vejan et al., 2016). The term "plant growth promoting bacteria" refers to those that colonise the roots of plants (rhizosphere) and enhance plant growth (Vejan et al., 2016). PGPR and their connections with plants are utilised commercially (Ahmad et al., 2008; Vejan et al., 2016) and hold great promise for sustainable agriculture. The inoculation of plants with beneficial microorganisms is a practice used in agriculture to increase plant growth and protect crops from different diseases and pests (Shaikh & Sayyed, 2015; Vejan et al., 2016). Examples of PGPR include strains of *Azospirillum brasilense*, *Bacillus amyloliquefaciens*, *Bacillus subtilis*, *Bradyrhizobium japonicum*, *Enterobacter cloacae*, *Gluconacetobacter diazotrophicus*, *Pantoea agglomerans*, *Pseudomonas fluorescens*, *Pseudomonas putida*,

Rhizobium leguminosarum, and *Sinorhizobium meliloti* (Bhattacharyya & Jha, 2017; Shaikh & Sayyed, 2015; Vejan et al., 2016).

Plant growth promoting-rhizobacteria are able to improve crop yield through a number of direct and indirect mechanisms (Castro et al., 2009; Shaikh & Sayyed, 2015; Vegan et al., 2016). Direct mechanisms includes nitrogen fixation, phytohormone production, phosphate solubilisation and increasing iron availability (Ahmad et al., 2008; Castro et al., 2009; Glick, 2014). Indirect mechanisms refer to bacterial traits that inhibit the functioning of one or more plant pathogenic organisms of fungi and/or bacteria. Indirect mechanisms include ACC deaminase, antibiosis, cell wall degrading enzymes, induced systemic resistance, quorum quenching and siderophores (Castro et al., 2009; Glick, 2014; Vegan et al., 2016). Applications of these mechanisms have been investigated in crops like maize, lettuce, wheat, oat, barley, pea, canola, soy, potato, tomato, lentil, radicchio, onion and cucumber (Shaikh & Sayyed, 2015; Vejan et al., 2016). The benefits of PGPR addition to plants can include an increase in seed germination rate, root growth, yield, leaf area, biocontrol, and chlorophyll content; tolerance to drought; and an increase in shoot and root weight (Castro et al., 2009; Shaikh & Sayyed, 2015; Vegan et al., 2016).

In 2014, Čolo et al. studied the effect of rhizosphere bacteria *Azotobacteria chroococcum*, *P. fluorescens* and *Bacillus subtilis* within the rhizosphere of onion production area. The author evaluated the ability of the strains to produce indole-acetic acid (IAA), siderophores and solubilisation of tricalcium phosphate (TCP). Their results showed that *Bacillus subtillis* was the most effective producer of IAA, whereas *Pseudomonas Fluorescens* strains were better at producing siderophore and solubilising phosphate (Čolo et al., 2014). Additionally, the best onion yield were observed in *B. subtillis* and *A. chroococcum* strains (Čolo et al., 2014). In 2015, Ernita et al. screened a complete of 136 rhizobacteria isolates were isolated from different rhizosphere soils in central areas of production of onions in Indonesia. This isolates were screened for their ability to strengthen growth and protect onions against bacterial blight disease-caused by *Xanthomonas axonopodis* pv. *allii* (Ernita et al., 2015). The results showed that all isolates produced indol-3-acetic acid with different concentrations (Ernita et al., 2015).

1.5 Bacterial communities in and on seed

Microbial communities in and on seeds may have useful or harmful effects on plant growth and fitness. Attachment of bacterial pathogens to the seed surface (coat) can enable the bacteria to evade the plant defence mechanisms within the seed, with the result that the seeds become a passive carrier of the pathogens (Robinson et al., 2016; Sharma & Kailash, 2014). The use of culture-independent methods to examine microbial communities in seeds can provide information on diversity of the bacterial communities on and within seed (Hardoim et al., 2012; Liu et al., 2012; Lopez-Velasco et al., 2013). Furthermore, culture-independent methods, such as 16S rRNA-based sequencing, has been used because of the ability to detect unculturable bacteria as well as bacteria that are in such low abundance or grow slowly compared to other bacteria that they could be missed by culture-dependent based protocols (Jackson et al., 2013; Qaisrani et al., 2019).

Research has indicated that soil type and irrigation methods influenced the structure of the seed endophyte community of maize (*Zea maydis*) and rice (*Oryzae sativ*a L.) (Barret et al., 2015; Hardoim et al., 2012; Johnston-Monji & Raizada, 2011). Furthermore, changes in seed microbiota structure can be described by the plant genetic features (e.g., seed size as well as geographic location of the production region, seed processing, harvesting methods and storage) (Aleklett & Hart, 2013; Barret et al., 2015). Buyer et al. (1999) reported similar results when the microbial composition of five seed types was assessed, viz., cucumber (*Cucumis sativus*), corn (*Z. mays*), soybean (*Glycine max*), sunflower (*Helianthus annuus*) and radish (*Raphanus sativus*), and their results showed that soil type affected the microbial community of spermosphere more than the seed type did.

In 2016, Khalaf & Raizada determined the taxonomic and functional diversity of seed-associated microbes of 21 cucurbit varieties belonging to seven species. They cultured 169 bacterial isolates belonging to the phyla Firmicutes and Proteobacteria. Furthermore, Adam et al. (2016) analysed the seed and rhizosphere microbiomes of 14 genotypes of oil seed pumpkin (*Cucurbita maxima*). Members of the Enterobacteriaceae, which included the genera *Erwinia* and *Pectobacterium*, dominated the seed microbiomes.

1.5.1 Methods used to describe microbial communities in and on seed

1.5.1.1 *Culture-dependent methods*

Culture-dependent methods to analyse bacterial communities are established on the capacity to culture bacteria on growth media. All bacteria require nutrients as well as growth factors that include amino acids, purines and pyrimidines to be able to grow in media (Agrios, 2005). Some bacteria also need vitamins for enzyme cofactors. The successful growth of bacteria on culture media depends on number of factors including nutritional requirements (C, N source), sensitivity to antibiotics, dye and other selective agents, and incubation temperature. (Agrios 2005; Prescott et al., 2005). Unfortunately, current culture-based methods have proven inadequate in terms of their ability to access all microbial life present in complex communities. For example, the study conducted by Lopez-Velasco et al. (2013) has shown that most isolates cultured from spinach (*Spinacia oleracea* L.) seed lots belonged to three phyla, Proteobacteria, Firmicutes and Actinobacteria, with the majority of the bacterial sequences unclassified.

The embryo and seed coat of various crops such as alfalfa (*Medicago sativa* L.), Norway spruce (*Picea abies*), cereals and cucurbits, produce a variety of culturable bacteria (Charkowski et al., 2002; Mundt & Hinkle, 1976; Lopez-Velasco et al., 2013). Methylotrophs and Rhizobia are normally correlated with seeds of bean (*Phaseolus vulgaris*) and soybean (*Glycine max* L) (Holland et al., 1992; Lopez-Velasco et al., 2013). Cultured bacterial populations of up to 3.5 x 10^5 CFU g^{-1} seed were present in rice (*Oryzae sativa* L.) seeds, and the genus *Stenotrophomonas maltophilia* and *Ochrobactrum* spp. were identified as the prominent bacterial genera (Hardoim et al., 2012). Ruiza et al. (2011) also isolated a number of genera including, *Curtobacterium*, *Microbacterium*, *Paenibacillus*, *Pantoea* and *Rhizobium* in fresh tissue of rice seeds.

1.5.1.2 *Culture-independent methods*

Although culture dependent methods of examining microbial communities remain important for complete characterisation of species, these methods have limitations when exploring the microbial world. Thus, culture independent methods are used as they can characterise microbes directly from the sample and reveal information related to the bacterial communities such as the diversity of microbes in different host (Schloss & Handelmans, 2004; Tringe et al., 2005). Current methods of culture independent profiling of bacterial communities typically rely on the amplification and sequencing of V4 region of the 16S rRNA gene (Caporaso et al., 2011; Kozick et al., 2013; Nelson, 2015).

Microbial profiling (also known as metagenomics) involves the sequencing of the 16S rRNA gene from an environmental sample, without prior culturing, in order to identify members of the bacterial community present in that sample (Handelsman et al., 1998; Nelson, 2015;). In principle, microbial profiling (metagenomics) is the study of microbial communities sampled directly from their natural environment, without prior culturing (Handelsman et al., 2007, Shah et al., 2011). Since over 99.8% of the microbes in some environments cannot be cultured, metagenomics offers a path to the study of their community structures (Shah et al., 2011; Streit & Schnitz, 2004). The communities can differ in the genera and species of specific microorganisms present, and they are usually defined in terms of diversity, identity, and abundance (Konopka, 2009; Links et al., 2014; Lopez-Velasco et al., 2013; Tringe & Hugenholtz, 2008). The 16S rRNA gene was the first gene used for microbial profiling and it is still widely used today (Barret et al., 2015; Link et al., 2014). The 16S rRNA gene is approximately 1.5 kb (*Escherichia coli*), highly conserved, encodes a structural RNA and forms part of the small ribosomal subunit (Brosius et al., 1978). The 16S rRNA gene contains nine hypervariable regions (V1-V9) that determine significant and differential sequence diversity among different bacteria (Chakravorty et al., 2007; Shah et al., 2011). Although it is still arguable which regions are best for species profiling, the amplification of 16S rRNA gene profiling target these hypervariable regions (Chakravorty et al., 2007; Shah et al., 2011). The limitation of 16S rRNA gene based analysis is that it is complicated by several artifacts, including chimeric sequences caused by PCR amplification and sequencing errors (Shah et al., 2011).

Various sequencing platforms such as 454 (Sogin et al., 2006), Ion Torrent (Junemann et al., 2012), PacBio (Fichot & Norman, 2013) and Illumina (Gloor et al., 2010), are used for surveying microbial communities. This review focuses on the Illumina platform as described Kozich et al. (2013). Estimated cycles from 300 to 500 of sequence data can be achieved on the Miseq platform and generate 8.5 Gb using paired 250 nt reads (Kozich et al., 2013). The first step in microbial profiling is genomic DNA extraction directly from the sample. A single PCR assay is used (Caporaso et al., 2011; Kozich et al., 2013). PCR amplification is performed with two primers that include an index sequence (only for the reverse primer), a 10-nt pad to avoid hairpin creation, the Illumina adapter sequence, a 2-nt linker that is non-complimentary to the 16S rRNA gene and a gene-specific primer (Kozich et al., 2013). The pad-linker-primer is used as the sequencing primer at the 5' end and the combined pad-linker-primer as the sequencing primer at the 3' end, to obtain a long read. The sequencing primer at the 3' end is a reverse complement of the combined pad-linker-primer to sequence the index region (Kozich et al.,

2013). With the 500 cycle reagents, this results in an index sequence and two 250-nt reads. Caporaso et al. (2011) published a collection of 2,168 reverse primers with different indices for the V4 region of the 16S rRNA gene. After that, Kozich et al. (2013) developed a dual index, paired-read approach that could be easily adjusted to other regions of the 16S rRNA gene. The Mothur software package (v. 1.30) is normally used to analyse Illumina sequence data (http://www.mothur.org/wiki/MiSeq_SOP) (Schloss et al., 2009).

1.6 Pathogenicity and virulence factors of seedborne plant pathogenic bacteria

Plant pathogenic bacteria have developed a number of different mechanisms that result in disease in the host plant (Buonaurio, 2008; Surico, 2013). Some of these mechanisms enable them to spread or to survive in their environments (Surico, 2013). The virulence factors or pathogenicity determinants used by bacteria to interact with the host can be unique to specific pathogens or shared among some plant pathogens (Buonaurio, 2008; Surico, 2013). For example, common mechanisms for adherence, invasion, evasion of host defences and damage to host cells are shared by different microbial pathogens (Buonaurio, 2008; Surico, 2013). However, most plant pathogens also have unique virulence factors that contribute to the pathogenic potential of a specific bacterium (Buonaurio, 2008; Surico, 2013).

Shurtleff & Averre (1997) defined pathogenicity as the ability of a pathogen to induce disease, whereas virulence was defined as the degree of pathogenicity of a given pathogen. Bacterial pathogens contain virulence genes that contributes in initiating disease symptoms in one or more plant hosts (Kannan et al., 2015; Melotto & Kunkel, 2013; Surico, 2013). Secretion systems play a crucial role in the virulence of plant pathogenic bacteria and strictly dependent on the presence of the particular secretion systems in the host cells (Buonaurio, 2008; Surico, 2013).

Pathogenicity and virulence genes are involved in the initiation of diseases in the host plant (Agrios, 2005; Surico, 2013). Kannan et al. (2015) and Surico (2013) reported that some of these genes are necessary for attachment of a pathogen to a plant surface, recognition of a host by a pathogen, penetration of the host and/or colonisation of host tissue. In the genus *Xanthomonas*, virulence factors such as the Type III secretion system (T3SS) and adhesion have been shown to be involved in the active transmission of the *Xanthomonas* spp. from plant to seeds (Darrasse et al., 2010; Darsonval et al., 2008, 2009). For example, attachment to seed is

a key step for the introduction of bacteria into the rhizosphere and would be advantageous for seedborne pathogen to attach to the host surface (Darrasse et al., 2010; Darsonval et al., 2009).

Pathogenic bacteria use multiple virulence factors acting together or individually at different stages of infection to cause disease in a susceptible host (Melotto & Kunkel, 2013; Surico, 2013). Hacker & Kaper (2000) reported that, in some plant pathogenic bacteria, virulence factors contribute to virulent, whereas in other non-pathogenic bacteria similar factors may be responsible for fitness and survival in the environment. In principle, a virulence factor is any molecule (in a secretion system, a plant cell-degrading enzyme, toxin production, a hormone, a siderophore, or an extracellular polysaccharide) can be cell-surface borne or can be secreted by bacterial cells or is transferred to an extracellular environment where it harms the host cells (Benali et al., 2014; Melotto & Kunkel, 2013; Surico, 2013). Virulence determinants or factors that play a significant role in the pathogenicity of plant pathogens are discussed below.

1.6.1 Secretion systems

Many virulence factors are secreted by bacterial plant pathogen to modulate host cell process from outside plant cells (Melotte & Kunkel, 2013; Prasannath, 2013). In pathogenic bacteria, secretion systems play a crucial role by producing surface structures for adhesion and bacterial movement, as well as to release cell wall-degrading enzymes, toxins, and plant pathogenic bacteria use dedicated protein secretion systems to secrete virulence factors from the cytosol of the bacteria into host cells or the host environment (Benali et al., 2014; Chang et al., 2014). Six forms of secretion pathways are recognised based on the proteins formed (Chang et al., 2014; Desvaux et al., 2004). Type I and II pathways secrete proteins into the host intercellular spaces, whereas type III and IV systems deliver proteins or nucleic acids directly into the host plant cell (Chang et al., 2014; Ponciano et al., 2003; Prasannath, 2013).

Type I secretion system

The type I secretion system (T1SS) is involved in the export of various molecules from the cytoplasm to the outside of the cell (Toth et al., 2006). The T1SS contains three proteins that make up a continuous channel (Toth et al., 2006). The inner membrane ATP binding cassette (ABC) protein transporter consists of a specific outer membrane protein known as the outer membrane fusion protein (OMP) and the so called membrane fusion protein (MFP), which is connected to the inner membrane and spans periplasmic space and extends to the outer membrane (Toth et al., 2006). The T1SS is found in almost all phytopathogenic bacteria and is involved in the secretion of toxins such as cyclolysin, hemolysins and rhizobiocin (Hennecke

& Verma, 1991). Proteases and lipases from the soft rot pathogenic bacterium *Dickeya dadantii* (formerly known as *Erwinia chrysanthemi*) are examples of plant pathogen effectors secreted via the T1SS (Palacios et al., 2001; Toth et al., 2006).

Type II secretion system

The type II secretion system (T2SS) is known as the secretion-dependent pathway and involves a two-step process in which proteins are first translocated across the inner membrane by the Sec or Tat pathway and then, after what might be an extremely short period, are transported from the periplasm to the exterior by an outer membrane secretin (Chang et al., 2014; Pfeilmeier et al., 2016; Prasannath, 2013). Studies have shown that the T2SS shares many characteristics with the type IV pilus structure (Chang et al., 2014; Pfeilmeier et al., 2016; Prasannath, 2013). Cianciotto (2005) reported that the T2SS is common in the γ-Proteobacteria, which contains the family Enterobacteriaceae.

The T2SS is necessary for the pathogenesis of bacteria belonging to the genera *Erwinia*, *Dickeya*, *Pectobacterium*, *Xanthomonas*, and *Ralstonia* on plants (Chang et al., 2014; Prasannath, 2013; Ray et al., 2000; Szczesny, et al., 2010; Toth et al., 2006). *Xanthomonas* and *Ralstonia* have two T2SS per cell, which are used for delivery of virulence factors such as pectinolytic and cellulolytic enzymes outside the bacterium (Prasannath, 2013). Soft-rot pathogens rely on the T2SS to attack the plant cell wall by secreting high amounts and multiple types of plant cell wall-degrading enzymes that cause loss of cell wall integrity, which contributes to characteristic rotting symptoms (Charkowski et al., 2012; Kazemi-Pour et al., 2004). Endopectate lyases that cleave polygalacturonate are the main class of degradative enzymes (Hassan et al., 2013). *Dickeya dadantii*, for example, secretes at least eight to ten different pectate lyases (Hassan et al., 2013).

Type III secretion system

The type III secretion system (T3SS) is found in many Gram negative bacteria of plant (Chang et al., 2014; Desvaux et al., 2004; He 2004; Pfeilmeier et al., 2016; Prasannath, 2013; Shariati et al., 2017; Tampakaki, 2014). The T3SS forms a needle-like structure that injects bacterial virulence "effector" proteins into the host cells (Cornelis & Gijsegem, 2000; Desvaux et al., 2004). The T3SS is important to the pathogenicity of plant pathogens in the genera *Pseudomonas*, *Xanthomonas*, *Ralstonia*, *Erwinia*, and *Pantoea* (Alfano & Collmer, 2004; Pfeilmeier et al., 2016; Prasannath, 2013), which colonize the intercellular spaces (apoplast) of

plants. Furthermore, the T3SS is capable of eliciting plant cell death at some stage in pathogenesis (Alfano & Collmer, 2004).

The T3SS pathway is encoded by a hypersensitive response and pathogenicity (*hrp*) and hypersensitive response conserved genes (*hrc*) (Bogdanove et al., 1996; Chang et al., 2014; Cornelis & Gijsegem, 2000; Pfeilmeier et al., 2016). The Hrc proteins direct release of T3SS substrates across the bacterial envelope (Alfano & Collmer, 2004; Cornelis & Gijsegem, 2000; Diepold & Wagner, 2014). A subset of the Hrp proteins (only partially defined) are themselves secreted by the T3SS and direct the translocation of effectors through host cell barriers (Alfano & Collmer, 2004; Cornelis & Gijsegem, 2000; Diepold & Wagner, 2014). The term "effector" means the subgroup of T3SS-substrates that function mainly inside host cells, but some proteins secreted by the T3SS of animal pathogens appear to have various functions in promoting effector translocation and acting directly in host cells (Cornelis & Gijsegem, 2000). Some effectors are also avirulence (AVR) proteins that are recognised by related host resistance (R) proteins and activating the hypersensitive response (HR) (Alfano & Collmer, 2004; Toth et al., 2006).

Type IV secretion Systems

Type IV secretion systems (T4SS) transport DNAs and/or proteins through the bacteria membrane (Alvarez-Martinez & Christie, 2009; Lawley et al., 2003; Rivera-Calzada et al., 2013). The T4SS system make large multiprotein network comprising of 12 proteins called VirB1 to VirB11 and VirD4. Alvarez-Martinez & Christie (2009) and Rivera-Calzada et al. (2013) reported that proteins VirB7, VirB9 and VirB10 connect into a 1.07 MegaDalton membrane-spanning core complex (CC), around which all other components assemble. Additionally, this complex is made of two parts, the O-layer inserted in the outer membrane and the I-layer inserted in the inner membrane (Alvarez-Martinez & Christie, 2009; Rivera-Calzada et al., 2013).

The T4SS is subdivided into three main clusters recognised by their purpose that includes the uptake or release of DNA, conjugation and proteins secreted by bacteria into the host cell (Lawley et al., 2003; Rivera-Calzada et al., 2013; Waksman & Fronzes, 2010). Alvarez-Martinez & Christie (2009) stated that T4SS group that facilitates DNA release and uptake also contributes to the development of genome diversity. Conjugation systems facilitates the transfer of DNA to recipient cells in a contact-dependent manner (Lawley et al., 2003). The T4SS encourages genome plasticity in bacteria and, then enables a rapid adaptive immune system to

variations in environment (Lawley et al., 2013; Waksman & Fronzes, 2010; Rivera-Calzada et al., 2013). Wallden et al. (2010) reported that T4SS are responsible for the delivery of antibiotic-resistance genes amongst pathogenic bacteria (Wallden et al., 2010). Alvarez-Martinez & Christie (2009) and McCullen & Binns (2006) demonstrated that the plasmid pTi from *Agrobacterium tumefaciens* encodes the VirB/VirD proteins, which contribute to the distribution of an oncogenic DNA portion called T-DNA into plant cells.

The VirB/D clusters consist of 12 proteins known as VirB1-VirB11 and VirD4 that form a transport apparatus called an envelope-spanning multiprotein (Alvarez-Martinez & Christie, 2009; Lessl & Lanka, 1994; Rivera-Calzada et al., 2013; Waksman & Fronzes, 2010). The proteins, VirB11, VirD4 and VirB4 control the substrate translocation and the assembly of the secretion system and they are known as the cytoplasmic ATPases connected with the inner membrane (Alvarez-Martinez & Christie, 2009; Rivera-Calzada et al., 2013; Waksman & Fronzes, 2010; Wallden et al., 2012). The VirB2 and VirB5 proteins are well-known as minor or major pilins and they are combined with the extracellular pilus (Alvarez-Martinez & Christie, 2009; Rivera-Calzada et al., 2013; Waksman & Fronzes, 2010; Wallden et al., 2012). Three proteins (TraN/VirB7, TraO/VirB9 and TraF/VirB10) form a large central structure (the core complex) of the T4SS structure (Alvarez-Martinez & Christie, 2009; Fronzes et al., 2009; Lawley et al., 2003; Rivera-Calzada et al., 2013).

Type V secretion system

The type V secretion system (T5SS) is widely present among Gram negative bacteria (Benali et al., 2014). The T5SS translocation system is committed to transfer of a single specific polypeptide known as the passenger domain, in a two-step process (Benali et al., 2014; Moreira et al., 2004). The signal sequence can either remain on the bacterial surface or cleave and then be delivered in the extracellular environment, through translocation of the passenger domain (Benali et al., 2014; Moreira et al., 2004). In Gram negative bacteria, the virulence factors related to the T5SS passenger domain are involved in biofilm formation, adhesins, toxins, enzyme production and cytotoxic activity (Benali et al., 2014; Leo et al., 2012). Plant pathogenic bacteria such as *Dickeya dadantii*, *Xanthomonas* spp. and *Xylella fastidiosa* use the T5SS as pathogenicity determinants (Benali et al., 2014; Moreira et al., 2004; Tseng et al., 2009).

Type VI secretion system

The type VI secretion system (T6SS) is a tool used by Gram negative bacteria to inject effector proteins into recipient cells (Ho et al., 2014; Mougous et al., 2006). During the T6SS secretion process, an intracellular tube complex composed of hexameric rings of haemolysin co-regulated proteins (Hcp) covered with a trimer of a valine-glycine repeat protein G (VgrG) and a proline-alanine-alanine-arginine (PAAR) repeat-containing protein, which is further surrounded by a sheath made of VipA/VipB heterodimers (also known as TssB/TssC) (Ho et al., 2014; Records, 2011; Salomon et al., 2014). Studies have shown that multiple T6SSs can be encoded within a single bacterial genome and each T6SS can have many Hcp, VgrG or PAAR-repeat containing proteins (Boyer et al., 2009; Records, 2011).

Shyntum et al. (2015) reported that the T6SS is responsible for pathogenicity in *P. ananatis*. In addition, the authors also reported that the T6SS contributes to the bacterial fitness of *P. ananatis*, not virulence. The T6SS have additionally been involved in the secretion of antimicrobials synthesised by bacteria, which may be particularly relevant for *P. ananatis* (Shyntum et al., 2015). Russell et al. (2014) reported the ability of T6SSs to transport effective antimicrobials straight into Gram negative pathogens makes the system attractive for the production of novel antimicrobial agents.

1.6.2 Phytotoxins produced by *Pseudomonas* species

Toxins play a significant role in virulence of several plant pathogenic bacteria (Agrios, 2005; Bender et al., 1999). *Pseudomonas* spp. produce phytotoxins that induce necrotic or chlorotic symptoms on the plant (Buonaurio, 2008). Syringomycin is a major virulence factor in *P. syringae*; it facilitates necrosis in the plasma membrane of the host plant (Agrios, 2005; Bender et al., 1999). Toth et al. (2006) stated that coronatine is initiated by conjugation of the polyketide coronafacic acid to coronamic acid and contributes to the virulence of *P. syringae* in the host plant (Toth et al., 2006). This virulence factor is required for initiation of disease symptoms in the plant by preventing the stomatal immune defence response (Bender et al., 1999). Benali et al. (2014) reported that coronatine prevents the pathogen-associated molecule pattern-triggered stomatal closure in *P. syringae* and *X. campestris.*

1.6.3 Quorum sensing and biofilm production

Cell to cell communication in bacteria that controls the mass of microbial inhabit using gene expression in response to the environment and chemical sensing system is called quorum sensing (Benali et al., 2014; Kanda et al., 2011; Melotto & Kunkel, 2013). The signalling molecules are produced during specific bacterial stages to promote physiological functions such as epiphytic growth, competition or colonisation and virulence (Kanda et al., 2011). In many plant pathogenic bacteria, the quorum sensing signal N-acyl homoserine lactones control factors such as enzyme production and exopolysaccharides (Teplitski et al., 2000). *Pantoea ananatis* and *P. agglomerans* use quorum sensing to cope with different environmental stress and also contributes to virulence of these bacteria on their host plants (Benali et al., 2014; Morohoshi et al., 2007; Sibanda et al., 2016).

Biofilm is defined as a complex multilayer cellular structure that attaches to host cells and is embedded within an exopolysaccharide (Benali et al., 2014). Melotto and Kunkel (2013) reported that biofilm formation may contribute to the early stages of tissue colonisation, e.g., by providing protection against antimicrobial compounds or by encouraging epiphytic survival within plant tissues. In addition, biofilms provide a defence mechanism for bacterial cells from extreme environmental conditions (Benali et al., 2014; Melotto & Kunkel, 2013). Benali et al. (2014) and Dow et al. (2003) stated that biofilms defend bacterial cells from host immune responses and antimicrobial responses. Plant pathogenic bacterial strains of *X. campestris* and *P. syringae* have been reported to produce biofilm, which contribute to the pathogenicity of the bacteria to plants (Dow et al., 2003; Keith et al., 2003).

1.6.4 Pili, fimbriae, and flagella

Bacteria depend on motility in order to spread rapidly between and within host plants (Meng et al., 2011; Weller-Stuart et al., 2017). Motility is an essential factor in the infection cycle of most plant pathogenic bacteria (Demir et al., 2011; Ichinose et al., 2013; Weller-Stuart et al., 2017). Additionally, motility enables the bacterial cells to detect appropriate points of attachment that allows bacteria to respond to changing concentrations of attractants and deterrents (Demir et al., 2011; Ichinose et al., 2013; Weller-Stuart et al., 2017). Flagella play a significant role in swimming motility, biofilm formation, aid in host attachment (Haiko & Westerlund-Wikström, 2013; Kang et al., 2002; Meng et al., 2011) and virulence (Haiko & Westerlund-Wikström, 2013; Shen & Ronald, 2002). Swimming motility was reported to

contribute in biofilm formation and colonisation of plant tissues in the vascular plant pathogens such as *R. solanacearum*, *P. stewartii* and *D. dadantii* (Jahn et al., 2008; Herrera et al., 2008).

Pili, fimbriae, and flagella are proteinaceous polymeric appendages that act as bacterial surface organelles (Bogino et al., 2013). Their roles include inter-bacterial interactions, bacterial-host interactions, facilitating motility and surface colonisation (Bogino et al., 2013). Pili and flagella are required for effective colonisation and contribute to virulence in plant, animal and human hosts (Bogino et al., 2013). The pili composition is an important element of biofilm formation and bacterial adhesion in pathogenic bacteria such as *Acidovorax citrulli* and *Xylella. fastidiosa* (Bahar et al., 2010). *Acidovorax citrulli* and *X. fastidiosa* uses Type IV pili (Tfp) and Type I pili, respectively, to colonise the host plant and induce symptoms (disease) (Bahar et al., 2010; Bogino et al., 2013).

1.7 Additional factors associated with pathogenicity

Adhesions and exopolysaccharides

Adhesions are glycoproteins and proteins that enable binding of bacteria to the host cell (Benali et al., 2014; Katzen et al., 1998). Adherence enables attachment of the bacteria and colonisation of the foliage or root tissues of the host and it plays a significant role in interaction between the pathogen and the plant (Alfano & Collmer, 2004; Benali et al., 2014; Bogino et al., 2013). Agrios (2005) and Cao et al. (2001) reported adhesion mechanism is not needed by some bacteria except when they are moving through the xylem and phloem. However, *A. tumefaciens* requires attachment to the plant surface as the first step in the transport of T-DNA into the host cell and, thereafter, to incite disease symptoms (Agrios, 2005; Cao et al., 2001; Prasannath, 2013).

Exopolysaccharides (EPSs) are carbohydrate polymers produced by bacteria, either firmly surrounding the cell or creating an extracellular slime (Benali et al., 2014; Kunkel & Chen, 2006; Melotto & Kunkel, 2013). EPSs produced by some plant pathogenic bacteria, for example, *X. campestris and E. amylovora,* enhance their virulence to the host plants (Denny, 1995; Dunger et al., 2007; Melotto & Kunkel, 2013; Prasannath, 2013). It was reported that EPSs protect bacterial cells growing in plant tissues from environmental stresses and toxins by enabling absorption of nutrients and water (Denny, 1995; Melotto & Kunkel, 2013; Prasannath; 2013). *Pseudomonas syringae* pv. *phaseolicola,* and *P. syringae* pv. *lachrymans* produce several EPSs such as L-glucuronic acid and alginate, which contribute to the symptom

expression in host plants (Fett & Dunn, 1989; Osman et al., 1986; Melotto & Kunkel, 2013). Virulence of *R. solanacearum* is enhanced by its ability to produce glutinous EPS while colonising vascular tissues (Bogino et al., 2013; Denny 1995; Melotto & Kunkel, 2013). It was reported that production of large amount of EPSs by bacteria colonising the vascular tissue, disrupt the transfer of water and nutrients within infected plant resulting in wilting of plant (Denny & Baek, 1991; Kao et al., 1992; Kunkel & Chen, 2006; Melotto & Kunkel, 2013).

Cell wall degrading enzymes

Cell wall degrading enzymes are responsible for pathogenicity in the host plant, thereby enabling cell penetration and tissue colonisation (Prasannath, 2013). Prasannath (2013) stated that these enzymes are required for symptom development in the host. Plant cell walls consist of three main polysaccharides: cellulose, hemicellulose and pectins (Denny, 1995; Prasannath, 2013). Plant cell wall degrading enzymes that break down pectinase, are an essential virulence component of the soft rot members of the Enterobacteriaceae (Toth et al., 2006). They are also important in other plant pathogens, e.g. *X. campestris* and *P. syringae* (Bauer & Collmer, 1997; van Sluys et al., 2002).

Iron acquisition

Iron is an important element for pathogenic bacteria (Benali et al., 20914; Buyer & Leong, 1986; Chu et al., 2010; Expert, 1999). It helps in several processes such as oxygen binding, redox response and as a cofactor for essential enzymes (Benali et al., 2014; Buyer & Leong, 1986). Studies have shown that plant pathogenic bacteria secrete siderophores, which compete with plant siderophores to bind iron released from damaged plant cells, and provide iron to the bacterial cells (Benali et al., 2013; Chu et al., 2010; Dellagi et al., 2009; Expert, 1999). Siderophores have been shown to play a major role as virulence factors for numerous plant pathogenic bacteria e.g., chrysobactin which is a catechol is secreted by *E. chrysanthemi* and *E. carotovora* (Alfano & Collmer, 2004; Benali et al., 2014).

1.8 Bacterial avirulence genes

Avirulence genes are classified by similar host plant resistance (R) genes, for which the resistance response is followed by a hypersensitive reaction (HR), a form of programmed cell death resulting from a burst of superoxide production and the expression of plant defence genes (Alfano & Collmer, 2004; Prasannath, 2013; Surico, 2013). Prasannath (2013) reported that avirulence genes cause a plant pathogen or pest to induce a resistance response in a host plant.

Many avirulence genes play a significant role as mediators of the interaction between plant pathogens and their hosts (Alfano & Collmer, 2004; Prasannath, 2013). This concept of avirulence genes as virulence factors was significantly facilitated by studies of the *Hrp* (HR and pathogenicity) pathway (Alfano & Collmer, 2004). Both animal and plant T3SS have been widely reviewed (Alfano & Collmer, 2004), and are noted for their capacity to distribute virulence and avirulence proteins into the host cells (Prasannath, 2013).

1.9 Whole genome sequencing

Recent developments in the field of genomics has led to the whole genome sequencing of a number of plant pathogens (van Sluys et al., 2002). One means of understanding how a pathogen causes disease is by sequencing its genome and mining the genome sequence for candidate genes involved in the pathogen interaction with the host (Vinatzer & Yan, 2008). DNA extraction from a pure culture (bacterial colony) is the first step when sequencing the genome. Several DNA extraction kits are available and considerations are based on which sequencing technology will be used. CTAB (cetyl) trimethylamonium bromide) is highly recommended (Porebski et al., 1997; Wilson, 1989). The method helps to differentiate extracted DNA from solutions containing high levels of polysaccharides (Porebski et al., 1997). The extracted DNA is then measured After DNA extraction, genomic DNA is measured with a Qubit® 2.0 fluorometer using the Qubit® dsDNA HS assay (Life Technologies, Carlsbad, US-CA) to determine the quality of the DNA (Huptas et al., 2016).

Library preparation is the first step of next generation sequencing, this step prepares DNA samples to be compatible with a sequencer (Del Angel et al., 2018; Ekblom & Wolf, 2014; Huptas et al., 2016). Library preparation is essential to the success of your next generation sequencing workflow (Huptas et al., 2016). It allows DNA to stick to the sequencing flow cell and allows the samples to be identified (Hupta et al., 2016). The methods used for library preparation are dependent on the proposed sequencing platform. The kits suggested by Illumina for library preparation from bacterial DNA are the Nextera DNA Flex (https://emea.support.illumina.com/downloads/nextera-dna-flex-library-prep-reference-guide-1000000025416.html) and Nextera XT™ Library Preparation Kits (http://support.illumina.com/downloads/nextera_xt_sample_preparation_guide_15031942.html).

The choice of which sequencing platform to use is important, when starting a genome sequencing project. There is a clear movement from traditional Sanger sequencing (~1 kb sequence reads) and Roche 454 sequencing (up to 800 bp) towards short read technologies such as Illumina HiSeq (usually 150 bp) (Del Angel et al., 2018; Ekblom & Wolf, 2014; Sanger et al., 1977; Slatko et al., 2018; Sohn & Nam, 2018). Several technologies offering this long read, such as Pacific Biosciences (up to 5 kb), Ion Torrent (~500 bp) and Illumina Moleculo (up to 10 kb), are also in the market (Edwards & Holts, 2013; Ekblom & Wolf, 2014; Slatko et al., 2018; Sohn & Nam, 2018). It was reported that, novel high-throughput sequencing technologies such as the Ion Torrent and Illumina overcome the time consuming requirement of generating libraries (Edwards & Holts, 2013; Slatko et al., 2018; Sohn & Nam, 2018). In the Illumina sequencing workflow, the adapters contain complementary sequences that allow the DNA fragments to bind to the flow cell (Del Angel et al., 2018, Hupta et al., 2016; Sohn & Nam, 2018).

After sequencing is completed, read trimming and filtering is performed using the next generation sequence quality control Toolkit (v2.2.3) with automatic detection of FASTQ irregular allowing for adapter-contaminated read removal (Huptas et al., 2016; Patel & Jain, 2017). FastQC is a tool that run either from the command line or through an interactive graphical user interface (www.bioinformatics.babraham.ac.uk/projects/fastqc). FastQC (v0.10.1) is used for visual confirmation of high quality (trimmed and filtered) read pairs (Huptas et al., 2016; www.bioinformatics.babraham.ac.uk/projects/fastqc). Additionally, it aims to provide a quality control report which can spot problems which originate either in the sequencer or in the starting library material (Huptas et al., 2016; Patel & Jain, 2017). Fastqc produces plots and statistics showing the average and range of the sequence quality values across the reads (Edward & Holt, 2013; www.bioinformatics.babraham.ac.uk/projects/fastqc). Reads losing their forward or reverse matching part during filtering are discarded from further analysis (Huptas et al., 2016).

Genome assembly

Genome assembly is defined as the process of grouping reads into contigs and then contigs into Scaffolds (Del Angel et al., 2018; Edward & Holt, 2013; Huptas et al., 2016; Sohn & Nam, 2018). A contig refers to a set of overlapping DNA segments that together represent a consensus region of DNA (Del Angel et al., 2018; Edward & Holt, 2013; Huptas et al., 2016; Sohn & Nam, 2018). A scaffold represents a large discontinuous region of DNA that comprises a

sequence of contigs and the gaps between them (Del Angel et al., 2018; Edward & Holt, 2013; Huptas et al., 2016; Sohn & Nam, 2018).

In scaffolding, assembled contigs are stitched together based on information from paired short reads (Del Angel et al., 2018; Edward & Holt, 2013). Determining if the assembly is ready for annotation is a key step towards successful genome annotation (Edward & Holt, 2013; Del Angel et al., 2018). N50 is frequently used as a standard metric to evaluate an assembly (Ekblom & Wolf, 2014; English et al., 2012). N50 is the shortest contig length needed to cover 50% of the genome, such that the sum of contig lengths covers 50% of the total size of all contigs (English et al., 2012). Assembly evaluation tools, such as Quast, compare the metrics between assemblies, and allow the user to make educated choices to further improve and select the best assembly (Gurevich et al., 2013).

Assembly tools, such as SPAdes, work best with smaller amounts of data and is well improved for bacterial projects (Bankevich et al., 2012; http://cab.spbu.ru/software/spades/). SPAdes makes use of paired and multi-sized de Bruijn graphs, whereas ABySS and Velvet are typical de Bruijn graph assemblers (Bankevich et al., 2012; Simpson et al., 2009; Zerbino et al., 2008). SPAdes is one of Eulerian de Bruijn graph assemblers, and is considered for single-cell sequencing (Bankevich et al., 2012; Huptas et al., 2016). This program represents the most stylish tool, which joins features for read error correction and contig mismatch correction prior to and after assembly (Bankevich et al., 2012; Huptas et al., 2016; Sohn & Nam, 2018). The *k*-mers from DNA fragment reads build the inner de Bruijn graph, which is used for contig assembly (Bankevich et al., 2012; Huptas et al., 2016; Sohn & Nam, 2018).

Genome annotation

Annotation is the process of finding genes on the genome sequence that includes the detection of ribosomal and transfer RNAs (rRNA and tRNA, respectively) encoded in the genome (Aziz et al., 2008; Edwards & Holt, 2013; http://rast.nmpdr.org). This step is performed once the ordered contigs have been finalised. Bacterial genome annotation is completed by uploading a genome assembly to an automated web-based tool such as Rapid Annotation using Subsystem Technology (RAST) v2.0 server (Aziz et al., 2008; http://rast.nmpdr.org). To annotated the genomes with RAST server, registration is required for genome submission and viewing of results. Users are directed to the "Jobs Overview once logged onto the server. To start a new job, upload the genome from the navigation bar, provide a valid taxonomy id, the organism's genus, species, and strain, as well as a nucleotide sequence file in FASTA format

(http://rast.nmpdr.org). After the annotation is complete, download the annotated genome in a variety of export formats (e.g. GenBank, FASTA, Excel) or browse the genome in the comparative environment of the SEED-Viewer (http://rast.nmpdr.org). The organism overview page contains basic information on the genome such as taxonomy, size, the number of contigs, the number of coding sequences and RNAs and counts of non-hypothetical and hypothetical gene annotations (http://rast.nmpdr.org). In addition, overview page contains the number of subsystems (*FIGfams*) that were automatically determined to be present in the genome (Aziz et al., 2008; http://rast.nmpdr.org).

Comparative genomics

Comparative genomics is the analysis of similarities and differences in the genome sequences and resulting features of related bacterial strains or species (Edwards & Holt, 2013; Touchman, 2010). Comparative genomics provides a powerful tool for studying evolutionary changes among organisms, helping to identify genes that are conserved or common among species as well as genes that give each organism unique characteristics (Edwards & Holt, 2013; Touchman, 2010). Genome alignment, synteny plots, and core and accessory genome clarification are basic tools for genomic studies (Bentley & Parkhill, 2004; Binnewies et al., 2006). Software such as BLAST Ring Image Generator (BRIG) (Alikhan et al., 2011; http://brig.sourceforge.net/), multiple alignment of conserved genomic sequences with rearrangements (Mauve) (Darling et al., 2010) are used to visualise of the genomes. The feature.txt file (Excel) downloaded from RAST server contains features such as the RAST ID, the location string, the feature type, the functional assignment, any alternated IDs found, and (for protein-coding genes) that are used for genome analyses (Aziz et al., 2008; http://rast.nmpdr.org).

The availability of sequenced genomes of different plant pathogenic bacteria has allowed for the discovery and comparison of genetic factors that contribute to the ability of certain microorganisms to prosper in different environments (Walterson & Stavrinides, 2015). *Pseudomonas syringae* genomes served as a starting point in determining virulence and host-specificity determinants using genomic sequences in bacterial plant pathogens (Lindeberg, 2012; Lindeberg et al., 2008). Dudnik & Dudler (2015) used comparative genomics to identify putative virulence-associated genes and other Poaceae-specific adaptations in several newly available genome sequences of *P. syringae* isolated from grass species. In their results, all strains possessed a small number of known T3SS effectors, highlighting the importance of non-T3SS virulence factors in pathogenicity of isolates of this species (Dudnik & Dudler, 2015).

A recent genetic analysis has contributed to our understanding of the mechanisms underlying *P. ananatis* phytopathogenesis (Weller-Stuart et al., 2017). For example, *P. ananatis* strains encode up to three T6SSs (Type VI secretion systems), which may play a role in pathogenesis in both plant and animal hosts (De Maayer et al., 2011; Shyntum et al., 2014). Type VI secretion system (T6SS-1) and T6SS-2 appear to be universal among *P. ananatis* strains and are assumed to play a role in antibiosis, fitness and niche adaptation (Shyntum et al., 2014). In a study conducted by Shyntum et al. (2015), it was shown that the T6SS-1 plays a role in pathogenicity on onion seedlings, and is an important factor in intra- and inter-species bacterial competition.

1.10 Conclusion

Understanding the survival, multiplication, and seed-to seedling transmission of plant pathogenic bacteria to seeds is central to study their pathogenesis. A thorough understanding of the epidemiology of seedborne pathogens may provide insights related to the need to suppress disease development and spread, and to develop effective and sustainable disease management practices. Although studies concerning the structure of seed-associated microbial communities are scarce, we need to understand what impact complex microbial communities can have on the onset and severity of seedborne pathogen transmission. Genomic analyses can provide insights into evolutionary adaptation processes of plant pathogenic bacteria that are seedborne. Determination of the diversity and distribution of effector proteins and other virulence genes within and across plant pathogenic species, pathovars and strains will allow us to understand how pathogens adapt to specific host plants and tissues, the evolutionary pathways available to these organisms, and the possible future for effective biocontrol of seedborne bacterial plant pathogens.

1.11 References

Adam, E., Bernhart, M., Müller, H., Winkler, J. & Berg, G. 2016. The *Cucurbita pepo* seed microbiome: genotype-specific composition and implications for breeding. *Plant and Soil* 422:35-39.

Agarwal, G., Stumpf, S., Kvitko, B. & Dutta, B. 2019. Center Rot of Onion. *The Plant Health Instr*. DOI: DOI:10.1094.

Agrios, G.N. 2005. Plant Pathology, 5th Ed. Elsevier Academic Press, Amsterdam.

Ahmad, F., Ahmad, I. & Khan, M.S. 2008. Screening of free-living rhizospheric bacteria for their multiple plant growth promoting activities. *Microbiol Res* 163:173-181.

Aleklett, K. & Hart, M. 2013. The root microbiota-a fingerprint in the soil? *Plant Soil* 370: 671-686.

Alfano, J.R. & Collmer, A. 2004. Type III secretion system effector proteins: double agents in bacterial disease and plant defense. *Annu Rev Phytopathol* 42:385-414.

Alvarez-Martinez, C.E. & Christie, P.J. 2009. Biological diversity of prokaryotic type IV secretion systems. *Microbiol Mol Biol Rev* 73:775-808.

Alvarez, A.M., Buddenhagen, I.W., Buddenhagen, E.S. & Domen, H.Y. 1978. Bacterial blight of onion, a new disease caused by a *Xanthomonas* sp. *Phytopathol* 68:1132-1136.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A. & Zagnitko, O. 2008. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9:75.

Bahar, O., de la Fuente, L. & Burdman, S. 2010. Assessing adhesion, biofilm formation and motility of *Acidovorax citrulli* using microfluidic flow chambers. *FEMS Microbiol Lett* 312:33-39.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, S.A., Lesin, M.V. Nikolenko, I.S., Pham, S., Andrey, D. Prjibelski, D.A., Pyshkin, V.A., Sirotkin, V.A., Vyahhi, N., Tesler, G., Alekseyev, A.M. & Pevzner, A.P. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455-477.

Barret, M., Briand, M., Bonneau, S., Préveaux, A., Valière, S., Bouchez, O., Hunault, G., Simoneau, P. & Jacques, M.A. 2015. Emergence shapes the structure of the seed microbiota. *Appl Environ Microbiol* 81:1257-1266.

Bauer, D.W. & Collmer, A. 1997. Molecular cloning, characterization, and mutagenesis of a

pel gene from *Pseudomonas syringae* pv. *lachrymans* encoding a member of the *Erwinia chrysanthemi* pelADE family of pectate lyases. *Mol Plant-Microbe Interact* 10:369-379.

Benali, S., Mohamed, B. & Eddine, H.J. 2014. Virulence strategies of phytopathogenic bacteria and their role in the plant disease pathogenesis. *African J Microbiol Res* 8:2809-2815.

Bender, C., Alarcón-Chaidez, F. & Gross, D.C. 1999. *Pseudomonas syringae* phytotoxins: mode of action, regulation, and biosynthesis by peptide and polyketide synthetases. *Microbiol Mol Biol Rev* 63:266-292.

Bentley, S.D. & Parkhill, J. 2004. Comparative genomic structure of prokaryotes. *Annu Rev Genet* 38:771-792.

Berg, G. 2009. Plant-microbe interactions promoting plant growth and health: perspectives for controlled use of microorganisms in agriculture. *Appl Microbiol Biotechnol* 84:11-18.

Berge, O., Monteil, C.L., Bartoli, C., Chandeysson, C., Guilbaud, C., Sands, D.C. & Morris, C.E. 2014. A user's guide to a database of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex. *PLoS ONE* 9(9):e105547.

Bhattacharyya, P.N. & Jha, D.K. 2012. Plant growth-promoting rhizobacteria (PGPR): Emergence in agriculture. *World J Microbiol Biotechnol* 28:1327-1350.

Binnewies, T.T., Motro, Y., Hallin, P.F., Lund, O., Dunn, D., La, T., Hampson, D.J., Bellgard, M., Wassenaar, T.M. & Ussery, D.W. 2006. Ten years of bacterial genome sequencing: comparative genomics based discoveries. *Funct Integr Genomics* 6:165-185.

Black, L., Conn, K.E, Gabos, B., Kao, J., Lutton, J.S. & Rosenberger, S.A. 2012. Onion Disease Guide. *Seminis* 1-72.

Bogdanove, A.J., Beer, S.V., Bonas, U., Boucher, C.A., Collmer, A., Coplin, D.L., Cornelis, G.R., Huang, H.C., Hutcheson, S.W., Panopoulos, N.J. & Van Gijsegem, F. 1996. Unified nomenclature for broadly conserved hrp genes of phytopathogenic bacteria. *Mol Microbiol* 20:681-683.

Bogino, C.P., Oliva, M.M., Fernando, G. Sorroche, G.F. & Giordano, W. 2013. The role of bacterial biofilms and surface components in plant-bacterial associations. *Int J Mol Sci* 14:15838-15859.

Boyer, F.G., Berthod, F.J., Vandenbrouck, Y. & Attree, I. 2009. Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: What can be learned from available microbial genomic resources? *BMC Genomics* 10:104.

Brady, C., Venter, S., Cleenwerck, I., Vanncanneyt, M., Swings, J. & Coutinho, A.T. 2007. A FAFLP system for the improved identification of plant-pathogenic and plant-

associated species of the genus *Pantoea*. System. *Appl Microbiol* 30:413-317.

Brady, C., Cleenwerck, I., Venter, S., Vancanneyt, M., Swings, J. & Coutinho, A.T. 2008. Phylogeny and identification of *Pantoea* species associated with plants, humans, and the natural environment based on multilocus sequence analysis (MLSA). *System Appl Microbiol* 31:447-460.

Brady, C., Goszczynska, T., Venter, S., Cleenwerck, I., De Vos, P., Gitaitis, D.R. & Coutinho A.T. 2011. *Pantoea allii* sp. nov., isolated from onion plants and seed. *Int J Syst Evol Microbiol* 61:932-937.

Brewster, J.L. 2008. Onions and other vegetable alliums (No 15). CABI.

Brosius, J., Palmer, M.L., Kennedy, P.J. & Noller, H.F. 1978. Complete nucleotide sequence of a 16S ribosomal RNA gene from Escherichia coli. *Proc Natl Acad Sci USA* 75:4801-4805.

Bulgarelli, D., Schlaeppi, K., Spaepen, S., Ver Loren van Themaat, E. & Schulze-Lefert, P. 2013. Structure and functions of the bacterial microbiota of plants. *Annu Rev Plant Biol* 64:807-838.

Bull, C.T. & Koike, S. T. 2015. Practical benefits of knowing the enemy: Modern molecular tools for diagnosing the etiology of bacterial diseases and understanding the taxonomy and diversity of plant-pathogenic bacteria. *Annu Rev Phytopathol* 53:157-180.

Bulletin EPPO. 2016. Bulletin 46:4-7.

Buonaurio, R. 2008. Infection and plant defence responses during plant-bacterial interaction, In: Plant-Microbe Interactions, E. Ait Barka and C. Clement (eds.), Research Signpost, India. Pp. 169-197.

Buyer, J.S. & Leong, J. 1986. Iron transport mediated antagonism between plant growth promoting and plant deleterious *Pseudomonas* strains. *J Biol Chem* 261:791-794.

Buyer, J.S., Roberts, D.P. & Russek-Cohen, E. 1999. Microbial community structure and function in the spermosphere as affected by soil and seed type. *Can J Microbiol* 45:138-144.

Cao, H., Baldini, R.L. & Rahme, L.G. 2001. Common mechanisms for pathogens of plants and animals. *Annu Rev Phytopathol* 39:259-284.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., Fierer, N. & Knight, R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc. Natl. Acad. Sci. USA 108 (Suppl 1):S4516-S4522.

Carr, E.A., Bonasera, J.M., Zaid, A.M., Lorbeer, J.W. & Beer, S.V. 2010. First report of bulb disease of onion caused by *Pantoea ananatis* in New York. *Plant Dis* 97:916.

Castro, R.O., Cornejo, H.A.C., Rodriguez, L.M. & Bucio, J.L. 2009. The role of microbial signals in plant growth and development. *Plant Signal Behav* 4:701712.

Chakravorty, S., Helb, D., Burday, M., Connell, N. & Alland, D. 2007. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *J Microbiol Methods* 69:330-339.

Chang, J.H., Desveaux, D. & Creason, A.L. 2014. The ABCs and 123s of bacterial secretion systems in plant pathogenesis. *Annu Rev Phytopathol* 52:317-345.

Charkowski, A.O., Barak, J.D., Sarreal, C.Z. & Mandrell, R.E. 2002. Differences in growth of *Salmonella enterica* and *Escherichia coli* O157:H7 on alfalfa sprouts. *Appl Environ Microbiol* 68:3114-3120.

Charkowski, A., Blanco, C., Condemine, G., Expert. D., Franza. T., Hayes, C., Hugouvieux-Cotte-Pattat, N., López Solanilla, E., Low, D., Moleleki, L., Pirhonen, M., Pitman, A., Perna, N., Reverchon, S., Rodríguez Palenzuela, P., San Francisco, M., Toth. I., Tsuyumu, S., van der Waals, J., van der Wolf, J., Van Gijsegem, F., Yang, C.H. & Yedidia, I. 2012. The role of secretion systems and small molecules in soft-rot *Enterobacteriaceae* pathogenicity. *Annu Rev Phytopathol* 50:425-49.

Chu, B., Garcia-Herrero, A., Johanson, T., Krewulak, K., Lau, C., Peacock, R.S., Slavinskaya, Z. & Vogel, H. 2010. Siderophore uptake in bacteria and the battle for iron with the host; a bird's eye view. *Biometals* 23:601-611.

Cianciotto, N.P. 2005. Type II secretion: a protein secretion system for all seasons. *Trends Microbiol* 13:581-588.

Compant, S., Duffy, B., Nowak, J., Clément, C. & Barka, E.A. 2005. Use of plant growth-promoting bacteria for biocontrol of plant diseases: principles, mechanisms of action, and future prospects. *Appl Environ Microbiol* 71:4951-4959.

Čolo, J., Hajnal-Jafari, T., Đurić, S., Stamenov, D. & Hamidović, S. 2014. Plant Growth Promotion Rhizobacteria in Onion Production. *Polish J Microbiol* 63:83-88.

Cornelis, G.R. & Van Gijsegem, F. 2000. Assembly and function of type III secretory systems. *Annu Rev Microbiol* 54:735-774.

Cother, E.J., Reinke, R., Mckenzie, C., Lanoiselet, V.M. & Noble, D.H. 2004. An unusual stem necrosis of rice caused by *Pantoea ananas* and the first record of this pathogen on rice in Australia. *Austral Plant Pathol* 33:495-503.

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

Danesh, R.Y., Amirnia, R., Tajbakhsh, M., Najafi, S. & Ghiyasi, M. 2014. Management of

seed borne bacterial diseases in seed production plots. Turkey 5th Seed Congress, 673-676.

Darling, A.E., Mau, B. & Perna, N.T. 2010. Progressive Mauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS ONE* 5(6): e11147.

Darrasse, A., Darsonval, A., Boureau, T., Brisset, M-N., Durand, K. & Jacques, M-A. 2010. Transmission of plant-pathogenic bacteria by nonhost seeds without induction of an associated defense reaction at emergence. *Appl Environ Microbiol* 76:6787-6796.

Darsonval, A., Darrasse, A., Durand, K., Bureau, C., Cesbron, S. & Jacques, M.A. 2009. Adhesion and fitness in the bean phyllosphere and transmission to seed of *Xanthomonas fuscans* subsp. *fuscans*. *Mol Plant Microbe Interact* 22:747-757.

Darsonval, A., Darrasse, A., Meyer, D., Demarty, M., Durand, K., Bureau, C., Manceau, C. & Jacques, M-A. 2008. The Type III secretion system of *Xanthomonas fuscans* subsp. *fuscans* is involved in the phyllosphere colonization process and in transmission to seeds of susceptible beans. *Appl Environ Microbiol* 74:2669-2678.

Dellagi, A. Seond, D., Rigault, M., Fagard, M.Simon, G. Saindrenan, P. & Expert, D. 2009. Microbial sideropheres exert a subtle role on *Arabidosis* during infection by manipulating the immune response and iron status. *Plant Physiol* 33:173-193.

De Maayer, P., Venter, S.N., Kamber, T., Duffy, B., Coutinho, T.A. & Smits, T.H.M. 2011. Comparative genomics of the type VI secretion systems of *Pantoea* and *Erwinia* species reveals the presence of putative effector islands that may be translocated by the VgrG and Hcp proteins. *BMC Genomics* 12:576-590.

Demir, M., Douarche, C., Yoney, A., Libchaber, A. & Salman, H. 2011. Effects of population density and chemical environment on the behaviour of *Escherichia coli* in shallow temperature gradients. *Phys Biol* 8:063001.

Denny, T.P. 1995. Involvement of bacterial polysaccharides in plant pathogenesis. *Annu Rev Phytopathol* 33:173-197.

Denny, T.P. & Baek, S.R. 1991. Genetic evidence that extracellular polysaccharide is a virulence factor of *Pseudomonas solanacearum*. *Mol Plant Microbe Interact* 4:198-206.

Department of Agriculture, Forestry and Fisheries, 2015. A profile of the South African onion market value chain. Department of Agriculture, Pretoria, South Africa.

Desvaux, M., Parham, N.J., Scott-Tucker, A. & Henderson, I.R. 2004. The general secretory pathway: a general misnomer? *Trends Microbiol* 12:306-309.

Del Angel, V.D., Hjerde, E., Sterck, L., Capella-Gutierrez, S., Notredame, C., Pettersson, O.V.,

Amselem, J., Bouri, L., Bocs, S., Klopp, C., Gibrat, J-F., Vlasova, A., Leskosek, B.L., Soler, L., Binzer-Panchal, M. & Lantz, H. 2018. Ten steps to get started in Genome Assembly and Annotation [version 1; referees: 2 approved]. F1000Research, 7:148.

Diepold, A. & Wagner, S. 2014. Assembly of the bacterial type III secretion machinery. *FEMS Microbiol Rev* 38:802-822.

Dow, J.M., Crossman, L., Findlay, K., He, Y.Q., Feng, J.X. & Tang, J.L. 2003. Biofilm dispersal in *Xanthomonas campestris* is controlled by cell-cell signaling and is required for full virulence to plants. *Proc Natl Acad Sci USA* 100:10995-11000.

Dudnik, A. & Dudler, R. 2015. Genomics-based exploration of virulence determinants and host-specific adaptations of *Pseudomonas syringae* strains isolated from grasses. *Pathogens* 3:121-148.

Dunger, G., Relling, V.M., Tondo, M.L., Barreras, M., Ielpi, L., Orellano, E.G. & Ottado, J. 2007. Xanthan is not essential for pathogenicity in citrus canker but contributes to *Xanthomonas* epiphytic survival. *Arch Microbiol* 188:27-135.

du Toit, L.J. 2004. Management of diseases in seed crops. Encyclopedia of Plant and Crop Science, 675-677.

Dutta, B., Avci, U., Hahn, M.G. & Walcott, R.R. 2012. Location of *Acidovorax citrulli* in infested watermelon seeds is influenced by the pathway of bacterial ingress. *Phytopathol* 102:461-468.

Dutta, B., Gitaitis, R., Smith, S. & Langston Jr., D. 2014. Interactions of seedborne bacterial pathogens with host and non-host plants in relation to seed infestation and seedling transmission. *PLoS ONE* 9(6): e99215.

Edens, D.G., Gitaitis, R.D., Sanders, F.H. & Nischwitz, C. 2006. First report of *Pantoea agglomerans* causing a leaf blight and bulb rot of onions in Georgia. *Plant Dis* 90:1551.

Edwards, D.J. & Holt, K.E. 2013. Beginner's guide to comparative bacterial genome analysis using next-generation sequence data. *Micro Inform Exper* 3:2.

Ekblom, R. & Jochen, B.W. 2014. Wolf A field guide to whole-genome sequencing, assembly and annotation. *Evol Appl* 7:1752-4571.

English, A.C., Richards, S., Han, Y., Wang, M., Vee, V., Qu, J., Qin, X., Muzny, D.M., Reid, J.G., Worley, K.C. & Gibbs, R.A. 2012. Mind the gap: upgrading genomes with Pacific Biosciences RS long-read sequencing technology. *PLoS ONE* 7(11): e47768.

Ernita, M., Habazar, T. & Jamsari, N. 2015. Screening of rhizobacteria from onion rhizosphere can induce systemic resistance to bacterial leaf blight disease on onion plants. *Int J Agric Sci*, 81-89.

Expert, D. 1999. Withholding and exchanging iron: Interactions between *Erwinia* spp. and their

plant hosts. *Annu Rev Phytopathol* 37:307-334.

Fabrice, D., Yann, M. & Kémal, B. 2018. Onion (*Allium Cepa*) Production in Urban and Peri Urban Areas: Financial Performance and Importance of This Activity for Market Gardeners in Southern Benin. *Curr Inves Agri Curr Res* 3(2).

Fett, W.F. & Dunn, M.F. 1989. Exopolysaccharides produced by phytopathogenic *Pseudomonas syringae* pathovars in infected leaves of susceptible hosts. *Plant Physiol* 89:5-9.

Fichot, E.B. & Norman, R.S. 2013. Microbial phylogenetic profiling with the Pacific Biosciences sequencing platform. *Microbiome* 1:10.

Gavini, F., Mergaert, J., Beji, A., Mielcarek, C., Izard, D., Kersters, K. & De Ley, J. 1989. Transfer of *Enterobacter agglomerans* (Beijerinck 1888) Ewing and Fife 1972 to *Pantoea* gen. nov. as *Pantoea agglomerans* comb. nov. and description of *Pantoea dispersa* sp. nov. *Int J Syst Bacteriol* 39:337-345.

Gent, D.H., Schwartz, F.H., Ishimaru, A.C., Louws, F.J., Cramer, A.R. & Lawrence, C.B. 2004. Polyphasic characterization of *Xanthomonas* strains from onion. *Phytopathol* 94:184-195.

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

Gitaitis, R., MacDonald, G., Torrance, R., Hartley, R., Sumner, D.R., Gay, D.J. & Johnson III, W.C. 1998. Bacterial streak and bulb rot of sweet onion II. Epiphytic survival of *Pseudomonas viridiflava* in association with multiple weed hosts. *Plant Dis* 82:935-938.

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

Gitaitis, R.D. & Walcott R.R. 2007. The epidemiology and management of seedborne bacterial diseases. *Annu Rev Phytopathol* 45:371-397.

Glick, B.R. 2014. Bacteria with ACC deaminase can promote plant growth and help to feed the world. *Microbiol Res* 169: 30-39.

Gloor, G.B., Hummelen, R., Macklaim, J.M., Dickson, R.J., Fernandes, A.D., MacPhee, R. & Reid, G. 2010. Microbiome profiling by Illumina sequencing of combinatorial sequence-tagged PCR products. *PLoS ONE* 5(10): e15406.

Goszczynska, T. & Serfontein, J.J. 1998. Milk-Tween agar, a semi-selective medium for

isolation and differentiation of *Pseudomonas syringae* pv. *syringae*, *Pseudomonas syringae* pv. *phaseolicola* and *Xanthomonas axonopodis* pv. *phaseoli*. *J Microbiol Methods* 32:65-72.

Goszczynska, T., Moloto, V.M., Venter, S.N. & Coutinho, T.A. 2006. Isolation and identification of *Pantoea ananatis* from onion seed in South Africa. *Seed Sci Technol J* 34:655-668.

Goszczynska, T., Venter, S.N. & Coutinho, T.A. 2007. Isolation and identification of the causal agent of brown stalk rot, a new disease of corn in South Africa. *Plant Dis* 91:711-718.

Goto, M. 1972. Bacterial leaf blight of onion in Japan. *Plant Dis Rep* 56:490-493.

Gupta, A. & Kurmar, R. 2020. Management of seedborne Disease: An integrated approach. Seed-Borne Diseases of Agricultural Crops: Detection, Diagnosis & Management. *Springer*, pp 717-745.

Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics.* 29:1072-1075.

Hacker, J. & Kaper, J.B. 2000. Pathogenicity islands and the evolution of microbes. *Annu Rev Microbiol* 54:641-679.

Haiko, J. & Westerlund-Wikström, B. 2013. The role of the bacterial flagellum in adhesion and virulence. *Biol* 2:1242-1267.

Handelsman, J., Rondon, M.R., Brady, S.F., Clardy, J. & Goodman, R.M. 1998. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem Biol* 5:R245-R249.

Handelsman, J., Tiedje, J.M., Alvarez-Cohen, L., Ashburner, M., Cann, I.K.O., Delong, E.F. Doolittle, W.F. Fraser-Liggett, C.M., Godzik, A.; Gordon, J.I., Riley, M., Schmid, M.B. & Reid, A.H. 2007. The new science of metagenomics: Revealing the secrets of our microbial planet. *The National Academies Press*.

Hardoim, P.R., Hardoim, C.C.P., van Overbeek, L.S. & van Elsas, J.D. 2012. Dynamics of seedborne rice endophytes on early plant growth stages. *PLoS ONE* 7(2): e30438.

Harman, G.E. 1982. Mechanisms of seed infection and pathogenesis. *Phytopathol* 325-329.

Hassan, S., Shevchik, E.V., Robert, X. & Hugouvieux-Cotte-Pattat, N. 2013. PelN is a new pectate lyase of *Dickeya dadantii* with unusual characteristics. *J Bacteriol* 195: 2197-2206.

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

Hennecke, H. & Verma, D.P.S. 1991. Advances in Molecular Genetics of Plant-Microbe Interactions, Volume 1, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Herrera, C.M, Koutsoudis, M.D., Wang, X. & von Bodman, S.B. 2008. *Pantoea stewartii* subsp. *stewartii* exhibits surface motility, which is a critical aspect of Stewart's wilt disease development on maize. *Mol Plant Microbe Interact* 21:1359-1370.

Ho, B.T., Dong, T.G. & Mekalanos, J.J. 2014. A view to a kill: The bacterial type VI secretion system. *Cell Host Microbe* 15:9-21.

Holland, M., Stebbins, N. & Polacco, J.C. 1992. Evidence of interactions in nitrogen metabolism between soybean and phylloplane bacteria. *Plant Physiol* 99:30A.

Humeau, L., Roumagnac, P., Picard, Y., Robène-Soustrade, I., Chiroleu, F., Gagnevin, L. & Pruvost, O. 2006. Quantitative and molecular epidemiology of bacterial blight of onion in seed production fields. *Phytopathol* 96:1345-1354.

Huptas, C., Scherer, S. & Wenning, M. 2016. Optimized Illumina PCR-free library preparation for bacterial whole genome sequencing and analysis of factors influencing de novo assembly. *BMC Res Notes* 9:269.

Hwang, M.S.H., Morgan, R.L., Sarkar, S.F., Wang, P.W. & Guttman, D.S. 2005. Phylogenetic characterization of virulence and resistance phenotypes of *Pseudomonas syringae*. *Appl Environ Microbiol* 71:5182-5191.

Ichinose, Y., Taguchi, F. & Mukaihara, T. 2013 Pathogenicity and virulence factors of *Pseudomonas syringae*. *J Gen Plant Pathol* 79:285-296.

International Seed Testing Association (ISTA), 2007. International Rules for Seed Testing. Bassersdorf, Switzerland: ISTA.

Jackson, C.R., Randolph, K.C., Osborn, S.L. & Tyler, H.L. 2013. Culture dependent and independent analysis of bacterial communities associated with commercial salad leaf vegetables. *BMC Microbiol* 13:274.

Jahn, C.E., Willis, D.K. & Charkowski, A.O. 2008. The flagellar sigma factor fliA is required for *Dickeya dadantii* virulence. *Mol Plant Microbe Interact* 21:1431-1442.

Johnson, K.B. & Stockwell, V.O. 1998. Management of fire blight: A case study in microbial ecology. *Annu Rev Phytopathol* 36:227-248.

Johnston-Monje, D. & Raizada, M.N. 2011. Conservation and diversity of seed associated endophytes in *Zea* across boundaries of evolution, ethnography and ecology. *PLoS ONE* 6(6): e20396.

Junemann, S., Prior, K., Szczepanowski, R., Harks, I., Ehmke, B., Goesmann, A., Stoye, J. & Harmsen, D. 2012. Bacterial community shift in treated periodontitis patients revealed by Ion Torrent 16S rRNA gene amplicon sequencing. *PLoS ONE* 7(8): e41606.

Kadota, I., Uehara, K., Shinohara, H. & Nishiyama, K. 2000. Bacterial blight of Welsh onion:

a new disease caused by *Xanthomonas campestris* pv. *allii* pv. nov. *J Gen Plant Pathol* 66:310-315.

Kaluzna, M., Janse, J.D. & Young, J.M. 2012. Detection and identification methods and new tests as used and developed in the framework of COST 873 for bacteria pathogenic to stone fruits and nuts. *J Plant Pathol* 94:S1.117-S1.126.

Kanda, E., Tatsuta, T., Suzuki, T., Taguchi, F., Naito, K., Inagaki, Y., Toyoda, K., Shiraishi, T. & Ichinose Y. 2011. Two flagellar stators and their roles in motility and virulence in *Pseudomonas syringae* pv. *tabaci* 6605. *Mol Gen Genomic* 285:163-174.

Kang, Y., Liu, H., Genin, S., Schell, M.A. & Denny, T.P. 2002. *Ralstonia solanacearum* requires type 4 pili to adhere to multiple surfaces and for natural transformation and virulence. *Mol Microbiol* 2:427-437.

Kannan, R.V., Bastas, K.K. & Antony, R. 2015. Plant pathogenic bacteria: An Overview In: Sustainable approaches to controlling plant pathogenic bacteria, pp.1-16. CRC Press.

Kao, C.C., Barlow, E. & Sequeira, L. 1992. Extracellular polysaccharide is required for wild-type virulence of *Pseudomonas solanacearum*. *J Bacteriol* 174:1068-1071.

Katzen, F., Ferreiro, D.U., Oddo, C.G., Ielmini, M.V., Becker, A., Puhler, A. & Ielpi, L. 1998. *Xanthomonas campestris* pv. *campestris* gum mutants: effects on xanthan biosynthesis and plant virulence. *J Bacteriol* 180:1607-1617.

Kazemi-Pour, N., Condemine, G. & Hugouvieux-Cotte-Pattat, N. 2004. The secretome of the plant pathogenic bacterium *Erwinia chrysanthemi*. *Proteomics* 4:3177-3186.

Keith, R.C., Keith, LM., Hernandez-Guzmam, G., Uppalapati, S.R. & Bender, C.L. 2003. Alginate gene expression by *Pseudomonas syringae* pv. *tomato* DC3000 in host and non-host. *Plant Microbiol* 149:1127-1138.

Khalaf, E.M. & Raizada, N.M. 2016. Taxonomic and functional diversity of cultured seed associated microbes of the cucurbit family. *BMC Microbiol* 16:131.

Kim, J. & Choi, O. 2012. An outbreak of onion center rot caused by *Pantoea ananatis* in Korea. *Plant Dis* 96:1576.

Kim, H.J., Lee, J.H., Kang, B.R., Rong, X., McSpadden Gardener, B.B., Ji, H.J., Park, C.-S. & Kim, Y.C. 2012. Draft genome sequence of *Pantoea ananatis* B1- 9, a non-pathogenic plant growth-promoting bacterium. *J Bacteriol* 194:729.

King, E.O., Ward, M.K. & Raney, D.E. 1954. Two simple media for the demonstration of pyocyanin and fluorescein. *J Lab Clin Med* 44:301-307.

Koike, S.T., Barak, J.D., Henderson, D.M. & Gilbertson, R.L. 1999. Bacterial blight of leek: a new disease in California caused by *Pseudomonas syringae*. *Plant Dis* 83:165-170.

Konopka, K. 2009. What is microbial community ecology? *Internat Soc Microbial Ecology*

3:1223-1230.

Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K. & Schloss, P.D. 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* 79:5112-5120.

Kunkel, N.B. & Chen, Z. 2006. Virulence strategies of plant pathogenic bacteria. Prokaryotes 2:421-440.

Lelliot, R.A., Billing, E. & Hayward, A.C. 1966. A determinative scheme for the fluorescent plant pathogenic *Pseudomonas*. *J Appl Bacteriol* 29:470-489.

Leo, J.C., Grin, I. & Linke, D. 2012. Type V secretion: mechanism(s) of autotransporter through the bacterial outer membrane. *Philos Trans R Soc Lond B Biol Sci* 367:1088-1101.

Lessl, M. & Lanka, E. 1994. Common mechanisms in bacterial conjugation and Ti-mediated T-DNA transfer to plant cells. *Cell* 77:321-324.

Lindeberg, M. 2012. Genome-enabled perspectives on the composition, evolution, and expression of virulence determinants in bacterial plant pathogens. *Annu Rev Phytopathol* 50:111-132.

Lindeberg, M., Myers, C.R., Collmer, A. & Schneider, D.J. 2008. Roadmap to new virulence determinants in *Pseudomonas syringae*: insights from comparative genomics and genome organization. *Mol Plant Microbe Interact* 21:685-700.

Links, M.G., Demeke, T., Gräfenhan, T., Hill, J.E., Hemmingsen, S.M. & Dumonceaux, T.J. 2014. Simultaneous profiling of seed-associated bacteria and fungi reveals antagonistic interactions between microorganisms within a shared epiphytic microbiome on *Triticum* and *Brassica* seeds. *New Phytol* 202:542-53.

Liu, Y., Zuo, S., Liwen Xu, L., Zou, Y. & Song, W. 2012. Study on diversity of endophytic bacterial communities in seeds of hybrid maize and their parental lines. *Arch Microbiol* 194:1001-1012.

Lopez-Velasco, G., Carder, P.A., Welbaum, G.E. & Ponder, M.A. 2013. Diversity of the spinach (*Spinacia oleracea*) spermosphere and phyllosphere bacterial communities. *FEMS Microbiol Lett* 346:146-154.

Louws, F.J., Fulbright, D.W., Stephens, C.T. & de Bruijin, F.J. 1994. Specific genomic fingerprints of phytopathogenic *Xanthomonas* and *Pseudomonas* pathovars and strains generated with repetitive sequences and PCR. *Appl Environ Microbiol* 60:2286-2295.

Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., Dow, M.,

Verdier, V., Beer, S.V., Machado, M.A., Toth, I., Salmond, G. & Foster, G.D. 2012. Top 10 plant pathogenic bacteria in molecular plant pathology. *Mol Plant Pathol* 13:614-629.

Marcelletti, S. & Scortichini, M. 2014. Definition of plant pathogenic *Pseudomonas* genomospecies of the *Pseudomonas syringae* complex through multiple comparative approaches. *Phytopathol* 104:1274-1282.

McCullen, C.A. & Binns, A.N. 2006. *Agrobacterium tumefaciens* and plant cell interactions and activities required for interkingdom macromolecular transfer. *Annu Rev Cell Dev Biol* 22:101-127.

Medrano, E.G. & Bell, A.A. 2007. Role of *Pantoea agglomerans* in opportunistic bacterial seed and boll rot of cotton (*Gossypium hirsutum*) grown in the field. *J Appl Microbiol* 102:143-143.

Melotto, M. & Kunkel, B.N. 2013. Virulence strategies of plant pathogenic bacteria. In: The Prokaryotes - Prokaryotic Physiology and Biochemistry (Rosenberg, E., ed.), pp. 6175.

Meng, F., Yao, J. & Allen, C. 2011. A MotN mutant of *Ralstonia solanacearum* is hypermotile and has reduced virulence. *J Bacteriol* 193:2477-2486.

Morohoshi, T., Nakamura, Y., Yamazaki, G., Ishida, A., Norihiro, K. & Ikeda, T. 2007. The plant pathogen *Pantoea ananatis* produces N-Acylhomoserine lactone and causes center rot disease of onion by quorum sensing. *J Bacteriol* 189:8333-8338.

Moloto, V.M., Goszczynska, T., du Toit, L.J. & Coutinho, T.A. 2017. A new pathovar of *Pseudomonas syringae*, pathovar *allii*, isolated from onion plants exhibiting symptoms of blight. *Eur J Plant Pathol* 147: 591-603.

Moreira, L.M., de Souza, R.F., Almeida Jr., N.F., Setubal, J.A., Oliveira, J.C., Furlan, L.R., Ferro, J.A. & da Silva, A.C. 2004. Comparative genomic analysis of citrus-associated bacteria. *Annu Rev Phytopathol* 42:162-184.

Morris, C.E., Sands, D.C., Vinatzer, B.A., Glaux, C., Guilbaud, C., Buffiere, A., Yan, S., Domiguez, H. & Thompson, M.B. 2008. The life history of the plant pathogen *Pseudomonas syringae* is linked to the water cycle. *ISME* 2:321-334.

Mougous, J.D., Cuff, M.E., Raunser, S., Shen, A., Zhou, M., Gifford, C.A., Goodman, A.L., Joachimiak, G., Ordoñez, C.L., Lory, S., Walz, T., Joachimiak, A. & Mekalanos, J.J. 2006. A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. *Science* 312:1526-1530.

Mundt, J.O. & Hinkle, N.F. 1976. Bacteria within ovules and seeds. *Appl Environ Microbiol* 32:694-698.

Myung, I.S., Joa, J.H. & Shim, H.S. 2011. Bacterial leaf spot of onion caused by *Pseudomonas*

syringae pv. *porri*, a New Disease in Korea. *Plant Dis* 95:1311.

Nelson, E.K. 2015. Encyclopedia of Metagenomics. *Springer, USA*. pg. XXI, 727.

Noble, D.H., Cother, E.J, Hailstones, D.L, Flack, M., Oxspring, L. & Hal, B. 2006. Characterisation of *Pseudomonas syringae* strains associated with a leaf disease of leek in Australia. *Eur J Plant Pathol* 115:419-430.

Nunez, J.J., Gilbertson, R.L. Meng, X. & Davis, R.M. 2002. First report of *Xanthomonas* leaf blight of onion in California. *Plant Dis* 86:330.

Osman, S.F., Fett, W.F. & Fishman, M.L. 1986. Exopolysaccharides of the phytopathogen *Pseudomonas syringae* pv. *glycinea*. *J Bacteriol* 166:66-71.

Palacios, J.L., Zaror, I., Martinez, P., Uribe, F., Opazo, P., Socias, T., Gidekel. M. & Venegas, A. 2001. Subset of hybrid eukaryotic proteins is exported by the type I secretion system of *Erwinia chrysanthemi*. *J Bacteriol* 183:1346-1358.

Parkinson, N., Aritua, V., Heeney, J., Cowie, C., Bew, J. & Stead, D. 2007 Phylogenetic analysis of *Xanthomonas* species by comparison of partial gyrase B gene sequences. *Int J Syst Evol Microbiol* 57:2881-2887.

Pareek, S., Sagar, N.A., Sharma, S. & Kumar, V. 2017. Onion (*Allium cepa* L). Fruit and vegetable phytochemicals: Chemistry and Human Health, 2nd Edition p. 1-14.

Patel, R.K. & Jain, M. 2012. Next generation sequencing quality control toolkit: a toolkit for quality control of next generation sequencing data. *PLoS ONE*. 7(2): e30619.

Pereira, I.S. & Tebaldi, N.D. 2013. Occurrence of *Xanthomonas axonopodis* pv. *allii* causing onion bacterial blight in Minas Gerais State, Brazil. *Summa Phytopathol* 39:297.

Pfeilmeier, S., Caly, D.L. & Malone, J.G. 2016. Bacterial pathogenesis of plants: future challenges from a microbial perspective. *Mol Plant Pathol* 17:1298-1313.

Picard, Y., Roumagnac, P., Legrand, D., Humeau, L., Robène-Soustrade, I., Chiroleu, F., Gagnevin, L. & Pruvost, O. 2008. Polyphasic characterization of *Xanthomonas axonopodis* pv. *allii* associated with outbreaks of bacterial blight on three *Allium* species in the Mascarene archipelago. *Phytopathol* 98:919-925.

Ponciano, G., Ishihara, H., Leach, J.E. & Tsuyumu, S. 2003. Bacterial effectors in plant disease and defence: Key to durable resistance? *Plant Dis* 87:1272-1282.

Porebski, S., Bailey, L.G. & Baum, B.R. 1997. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Mol Biol Rep* 15:8.

Prasannath, K. 2013. Pathogenicity and virulence factors of phytobacteria. *Sch Acad J Biosci* 1:24-33.

Prescott, L.M., Harley, J.P. & Klein, D.A. 2005. Bacteria: The *Proteobacteria*. In:

Microbiology, 6th Edition pp. 482-483. New York: McGraw-Hill.

Qaisrani, M.M., Zaheer, A., Mirza, M.S., Naqqash, T., Qaisrani B.T, Hanif, M. K., Rasool, G., Abdulla, K., Saleem, M., Ullah Mohammad, U., Jamal, S., Mirza, Z., Karim, S. & Rasool, M. 2019. A comparative study of bacterial diversity based on culturableand culture-independent techniques in the rhizosphere of maize (*Zea mays* L.). *Saudi J Biol Sci* 26:1344-1351.

Randle, W.M., Kopsell, D.A., Kopsell, D.E., Snyder, R.L. & Torrance, R. 1998. Field sampling short day onions for bulb pungency. *Hort Technol* 3:329-332.

Ray, S.K., Rajeshwari, R. & Sonti, R.V. 2000. Mutants of *Xanthomonas oryzae* pv. *oryzae* deficient in general secretory pathway are virulence deficient and unable to secrete xylanase. *Mol Plant Microbe Interact* 13:394-401.

Robène, I., Perret, M., Jouen, E., Escalon, E., Maillot, N-V., Chabirand, A., Moreau, A., Laurent, A., Chiroleu, F. & Pruvost, O. 2015. Development and validation of a real-time quantitative PCR assay to detect *Xanthomonas axonopodis* pv. *allii* from onion seed. *J. Microbiol Methods s* 114:78-86.

Records, A.R. 2011. The type VI secretion system: A multipurpose delivery system with a phage-like machinery. *Mol Plant-Microbe Interact* 24:751-757.

Rivera-Calzada, A., Fronzes, R., Savva, C.G., Chandran, V., Lian, P.W., Laeremans, T., Pardon, E., Steyaert, J., Remaut, H., Waksman, G. & Orlova, E.V. 2013. Structure of a bacterial type IV secretion core complex at subnanometre resolution. *EMBO Journal* 32:1195-1204.

Robinson, J.R., Fraaije, A.B., Clark, M, I., Jackson, W.R., Hirsch, R.P. & Mauchline, H.T. 2016. Wheat seed embryo excision enables the creation of axenic seedlings and Koch's postulates testing of putative bacterial endophytes. *Sci Rep* 6:25581.

Roumagnac, P., Gagnevin, L. & Pruvost, O. 2000. Detection of *Xanthomonas* sp., the causal agent of onion bacterial blight, in onion seeds using a newly developed semi-selective isolation medium. *Eur J Plant Pathol* 106:867-877.

Roumagnac, P., Gagnevin, L., Gardan, L., Sutra, L., Manceau, C., Dickstein, E.R., Jones, J.B., Rott, P. & Pruvost, O. 2004. Polyphasic characterization of xanthomonads isolated from onion, garlic and Welsh onion (*Allium* spp.) and their relatedness to different *Xanthomonas* species. *Int J Syst Evol Microbiol* 54:15-24.

Ruiza, D., Agaras, B., de Werra, P., Wall, L.G. & Valverde, C. 2011. Characterization and screening of plant probiotic traits of bacteria isolated from rice seeds cultivated in Argentinia. *J Microbiol* 49:902-912.

Russell, A.B., Peterson, S.B. & Mougous, J.D. 2014. Type VI secretion effectors: poisons with

a purpose. *Nat Rev Microbiol* 12:137-148.

Salomon, D., Kinchb, L.N., Trudgianc, D.C., Guoc, X., Klimkoa, J.A. Grishinb, N.V., Mirzaeic, H. & Orth, K. 2014. Marker for type VI secretion system effectors. *Microbiol* 111: 9271-9276.

Samson, R., Shafik, H., Benjama, A. & Gardan, L. 1998. Description of the bacterium causing blight of leek as *Pseudomonas syringae* pv. (pv. nov.). *Phytopathol* 88:844-850.

SANSOR Annual Report 2018/2019. 2021. South African National Seed Organisation. Pretoria, South Africa.

SANSOR Annual Report 2013/2014. 2015. South African National Seed Organisation. Pretoria, South Africa.

Sanger, F., Nicklen, S. & Coulson, A.R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74:5463-5467.

Schaad, W.N., Mortensen, N.C., Li, J., Feng, J., Luo, L., Mazzaglia, A. & Balestra, M.G. 2014. Technical challenges for specific, sensitive detection of seedborne bacterial pathogens. *Plant Pathol* 6:59-66.

Schloss, P.D. & Handelsman, J. 2004. Status of the microbial census. *Microbiol Mol Biol Rev* 68:686-691.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R, Hartmann, M., Hollister, E.B., Lesniewski R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres. B, Thallinger, G.G., Van Horn, D.J. & Weber, C.F. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75:7537-7541.

Schmid, H., Schubert, S., Weber, C. & Bogner, J.R. 2003. Isolation of a *Pantoea dispersa*-like strain from a 71-year-old woman with acute myeloid leukemia and multiple myeloma. *Infection* 31:225-420.

Schwartz, H. & Gent, D.H. 2007. *Xanthomonas* Leaf Blight of Onion. Colorado State University.

Schwartz, H.F. & Otto, K. 2000. First report of a leaf blight of onion by *Xanthomonas campestris* in Colorado. *Plant Dis* 84:922.

Serfontein, J.J. 2001. *Xanthomonas* blight of onion in South Africa. *Plant Dis* 85:442.

Shah, N., Tang, H., Doak, G.T. & Ye, Y. 2011. Comparing bacterial communities inferred from 16S rRNA gene sequencing and shotgun metagenomics. *Pac Symp Biocomput* 165-176.

Shaikh, S.S. & Sayyed, R.Z. 2015. Role of Plant Growth-Promoting Rhizobacteria and Their bormulation in biocontrol of plant diseases. plant microbes symbiosis: *Applied Facets*

18:337-351.

Shariati, V., Malboobi, M.A., Tabrizi, Z., Tavakol, E., Owilia, P. & Safari, M. 2017. Comprehensive genomic analysis of a plant growth-promoting rhizobacterium *Pantoea agglomerans* strain P5. *Scientific Reports* 7:15610.

Sharma, D.K. & Kailash, A. 2014. Incidence and histopathological study of *Xanthomonas axonopodis* pv. *vesicatoria* in tomato (*Lycopersicon esculentum* Mill.) seeds. *Int J Agric Technol* 10:33-242.

Shen, Y. & Ronald, P. 2002. Molecular determinants of disease and resistance in interactions of *Xanthomonas oryzae* pv. *oryzae* and rice. *Microbes Infect* 4:1361-1367.

Shurtleff, M.C. & Averre, C.W. 1997. Glossary of plant-pathological terms, APS Press, St. Paul, Minnesota, pp 361.

Shyntum, D.Y., Venter, S.N., Moleleki, L.N., Toth, I. & Coutinho, T.A. 2014. Comparative genomics of type VI secretion systems in strains of *Pantoea ananatis* from different environments. *BMC Genomics* 15:163-178.

Shyntum, D.Y., Theron, J., Venter, S.N., Moleleki, L.N., Toth, I. & Coutinho, T.A. 2015. *Pantoea ananatis* utilizes a type VI secretion system for pathogenesis and bacterial competition. *Mol Plant-Microbe Interact* 28:420-431.

Sibanda, S., Theron, J., Shyntum, D.Y., Moleleki, L. & Coutinho, T.A. 2016. Characterization of two Lux I/R homologs in *Pantoea ananatis* LMG 2665T. *Can J Microbiol* 62:893-903.

Simpson, J.T., Wong, K, Jackman, S.D., Schein, J.E., Jones, S.J. & Birol, I. 2009. ABySS: a parallel assembler for short read sequence data. *Genome Res*. 19:1117-1123.

Slatko, B. E., Gardner, A.F. & Ausubel, M.F. 2018. Overview of Next Generation Sequencing Technologies. *Curr Protoc Mol Biol* 122(1): e59.

Sogin, M.L., Morrison, H.G., Huber, J.A., Welch, D.M., Huse, S.M., Neal, P.R, Arrieta, J.M. & Herndl, G.J. 2006. Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc Natl Acad Sci USA* 103:12115-12120.

Sohn, J-I. & Nam, J-W. 2018. The present and future of de novo whole-genome assembly. *Brief Bioinform*, 19:23-40.

Starke Ayres, 2015. Onion production guideline. South Africa.

Streit, W. & Schmitz, A. 2004. Metagenomics-the key to the uncultured microbes. *Curr Opin Microbiol* 7:492-498.

Surico, G. 2013. The concepts of plant pathogenicity, virulence/avirulence and effector proteins by a teacher of plant pathology. *Phytopathol Mediterr* 52:399-417.

Szczesny, R., Jordan, M., Schramm, C., Schulz, S., Cogez, V., Bonas, U. & Büttner, D. 2010.

Functional characterization of the Xcs and Xps type II secretion systems from the plant pathogenic bacterium *Xanthomonas campestris* pv. *vesicatoria*. *New Phytol* 187:983-1002.

Tampakaki, A.P. 2014. Commonalities and differences of T3SSs in rhizobia and plant pathogenic bacteria. *Front Plant Sci* 5:114.

Teplitski, M., Robinson, J.B. & Bauer, W.D. 2000. Plant secretes substances that mimic bacterial N-acyl homoserine lactone signal activities and affect population density-dependent behaviours in associated bacteria. *Phytopathol* 13:637-648.

Tseng, T.T., Tyler, B.M. & Setubal, J.C. 2009. Protein secretion systems in bacterial host systems in bacterial-host associations and their description in the gene ontology. *BMC Microbiol* 9 (Suppl 1):S2.

The National Agricultural Directory, 2011. Field crops and horticulture. Vegetables. Department of Agriculture, Pretoria, South Africa.

Toth, I.K., Pritchard, L. & Birch. P.R.J. 2006. Comparative genomics reveals what makes an enterobacterial plant pathogen. *Annu Rev Phytopathol* 44:305-336.

Touchman, J. 2010. Comparative genomics. *Nature Educ Knowledge* 3:13.

Tringe, S.G., von Mering, C., Kobayashi, A., Salamov, A.A., Chen, K., Chang, H.W., Podar, M., Short, J.M., Mathur, E.J., Detter J.C., Bork, P., Hugenholtz, P. & Rubin, E.M. 2005. Comparative metagenomics of microbial communities. *Science* 308:554-557.

Tringe, S.G. & Hugenholtz, P. 2008. A renaissance for the pioneering 16S rRNA gene. *Curr Opin Microbiol* 11:442-446.

Tsuji, M. & Takikawa, Y. 2018. *Pseudomonas syringae* pv. *alliifistulosi* pv. nov., the causal agent of bacterial leaf spot of onions. *J Gen Plant Pathol* 84:343-358.

van Sluys, M.A., Monteiro-Vitorello, C.B., Camargo, L.E.A., Menck, C.F.M. da Silva, A.C.R., Ferro, J.A., Oliveira, M.C., Setubal, J.C., Kitajima, J.P. &. Simpson, A.J. 2002. Comparative genomic analysis of plant-associated bacteria. *Annu Rev Phytopathol* 40:169-189.

Vejan, P., Abdullah, R., Khadiran, T., Ismail, S. & Boyce, A.N. 2016. Role of Plant Growth Promoting Rhizobacteria in Agricultural Sustainability - A review. *Molecules* 21:573.

Vinatzer, B.A. & Yan, S. 2008. Mining the genomes of plant pathogenic bacteria: how not to drown in gigabases of sequence. *Mol Plant Pathol* 9:105-118.

Waksman, G. & Fronzes R. 2010. Molecular architecture of bacterial type IV secretion systems. *Trends Biochem Sci* 35:691-698.

Walcott, R.R. 2003. Detection of seedborne pathogens. *Hort Technol* 13:40-47.

Walcott, R.R., Gitaitis, R.D., Castro, A.C., Sanders Jr., F.H. & Diaz-Perez, J.C., 2002. Natural

infestation of onion seed by *Pantoea ananatis*, causal agent of center rot. *Plant Dis* 86:106-111.

Wallden, K., Rivera-Calzada, A. & Waksman, G. 2010. Type IV secretion systems: versatility and diversity in function. *Cell Microbiol* 12:1203-1212.

Wallden, K., Williams, R., Yan, J., Lian, P.W., Wang, L., Thalassinos, K., Orlova, E.V. & Waksman, G. 2012. Structure of the VirB4 ATPase, alone and bound to the core complex of a type IV secretion system. *Proc Natl Acad Sci USA* 10:11348-11353.

Walterson, A.M. & Stavrinides, J. 2015. *Pantoea*: insights into a highly versatile and diverse genus within the Enterobacteriaceae. *FEMS Microbiol Rev* 968-84.

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

Weller-Stuart, T., Toth, I., de Maayer. P. & Coutinho, T. 2017. Swimming and twitching motility are essential for attachment and virulence of *Pantoea ananatis* in onion seedlings. *Mol Plant Pathol* 18:734-745.

Wilson, K. 1989. Preparation of genomic DNA from bacteria. In Current protocols in molecular biology, (eds. F.M. Ausubel, R. Brent, R.E. Kimgston, D.D. Moore, J.G. Seidman, J.A. Smith and K. Struhl), 241-245, John Wiley & Sons, Inc, New York, N.Y.

Zaid, A.M., Bonaser, J. M. & Beer, S.V. 2012. OEM-A new medium for rapid isolation of onion-pathogenic and onion-associated bacteria. *J Microbial Methods*, 19:520-526.

Zerbino, D.R. & Birney, E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res*. 18:821-829.

Figure 1.12

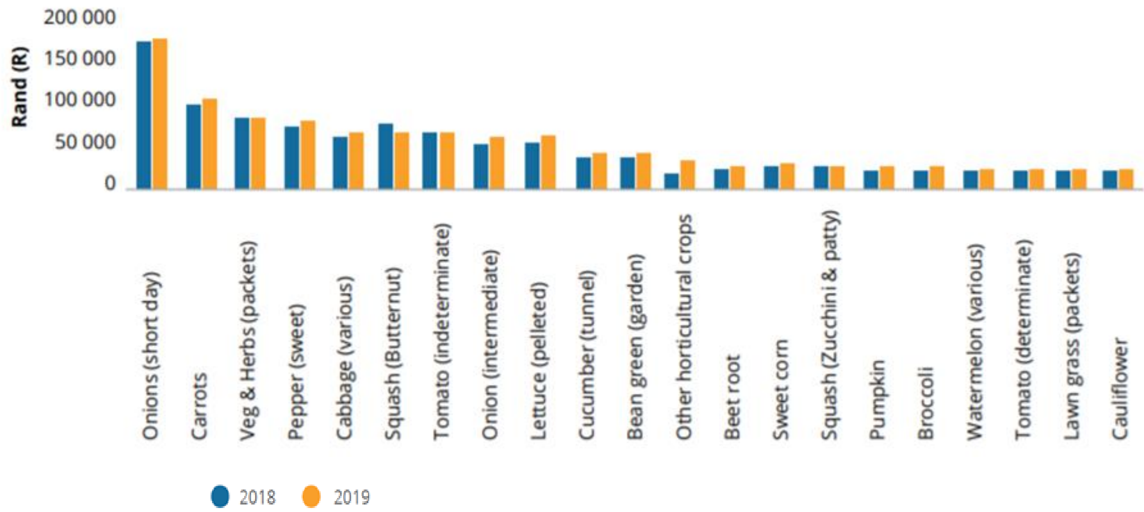


Figure 1.1 Top 20 vegetable seed types sold in South Africa: 2018 verse 2019 (SANSOR Annual Report, 2020)

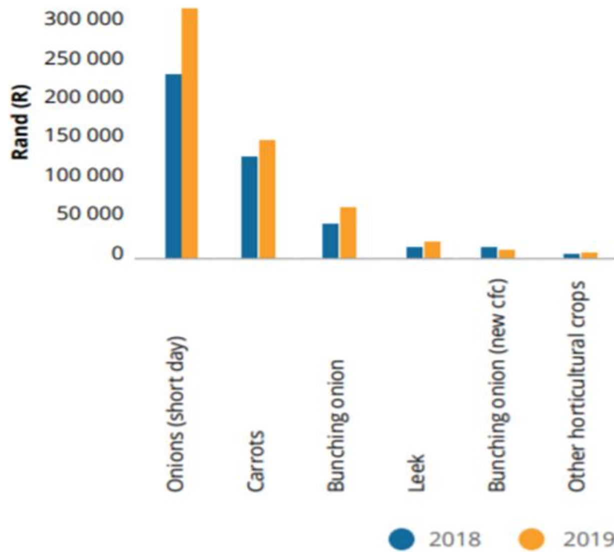


Figure 1.2 Top five seed production crops in South Africa 2018 verses 2019 (SANSOR Annual Report, 2020).

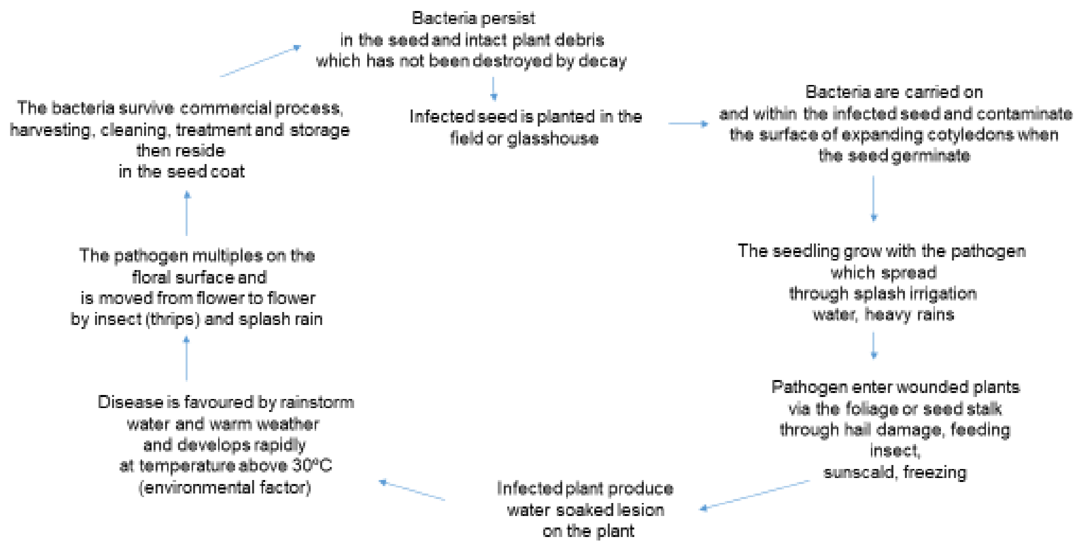


Figure 1.3 Disease cycle of a typical seedborne bacteria of onion.

Chapter 2

A new pathovar of *Pseudomonas syringae*, pathovar *allii,* isolated from onion plants exhibiting symptoms of blight

Published as:

V.M. Moloto, T. Goszczynska, L.J. du Toit, T.A Coutinho (2017) A new pathovar of *Pseudomonas syringae*, pathovar *allii,* isolated from onion plants exhibiting symptoms of blight. European Journal of Plant Pathology 147, 591-603.

2.1 Abstract

Bacterial pathogens of onion (*Allium cepa*) plants and their undetected presence in seed can cause substantial losses to onion producers. In this study, 23 *Pseudomonas syringae* strains were isolated from five onion plants and 18 onion seeds. The symptoms on leaves and seed stalks were irregular lesions with necrotic centres and water soaked margins. The aim of the study was to characterize these *P. syringae* strains using Biolog GN III carbon source utilization, multilocus sequence typing (MLST) based on partial sequences of four housekeeping genes (*cts, gapA*, *gyrB* and *rpoD*), and to determine whether or not the strains were pathogenic on onion (cv. Granex 33), chive (*Allium schoenoprasum* cv. Grasiue), leek (*Allium porrum* cv. Giant Italian) and spring onion (*Allium fistulosum* cv. Salotte) plants. Both Biolog analysis and MLST analysis separated onion strains into two clusters, one supporting the existence of a new pathovar of *P. syringae*, and the other corresponding to *P. syringae* pv. *porri*. *Pseudomonas syringae* strains belonging to the new pathovar we pathogenic only on onion plants of the *Allium* spp. tested. The results of this study revealed that bacterial blight of onion in South Africa is caused by two pathovars of *P. syringae sensu lato*, namely, the newly described pathovar, *allii*, and *P. syringae* pv. *porri*. The symptoms caused by these two pathovars in the field were indistinguishable.

Keywords Phenotypic amplification Cluster core genome

2.2 Introduction

Pseudomonas syringae (van Hall 1902) is a Gram negative bacterium of major agricultural and economical concern (Young 2010). This bacterium belongs to the Gammaproteobacteria, and includes all fluorescent pseudomonad strains that are oxidase negative, arginine dihydrolase negative and induce a hypersensitive reaction when inoculated into tobacco plants (Palleroni 1984). *Pseudomonas syringae sensu lato* is responsible for a variety of bacterial blight, speck and spot diseases on a wide range of important crop species. Although a large number of plant species can be infected by strains of *P. syringae*, every strain shows some degree of host specificity, i.e., induces symptoms only in some plant species (Bull and Koike 2015; Lamichhane et al. 2015).

Pseudomonas syringae sensu lato is subdivided taxonomically into more than 60 pathovars (pathogenic varieties) based on the host plant species from which the strain was isolated, and

includes pathovars from, for example, apple (*Malus* spp.), beet (*Beta vulgaris*), bean (*Phaseolus* spp.), brassicas (*Brassica* spp.), cucumber (*Cucumis sativus*), various ornamental flowering genera, oat (*Avena* spp.), olive (*Osmanthus* spp.), onion (*Allium cepa*), pea (*Pisum sativum*), tobacco (*Nicotiana* spp.), tomato (*Solanum lycopersicon*) and rice (*Oryza* spp.) (Young 2010). The term pathovar refers to a strain or set of strains with the same or similar pathological characteristics, differentiated at the infra-subspecific level from other strains of the same species or subspecies on the basis of distinctive pathogenicity to one or more plant hosts (Dye et al. 1980). *Pseudomonas syringae* pathovars have been defined previously both biochemically and by pathogenicity tests. The LOPAT determinative tests (levan, oxidase, potato rot, arginine dihydrolase and tobacco hypersensitive reaction) are used to differentiate plant pathogenic *Pseudomonas* spp., but the tests are not useful for differentiating pathovars within *P. syringae* (Bull et al. 2011; Lelliott et al. 1966). DNA-DNA hybridization has enabled reclassification of *P. syringae* species into nine different genomospecies, which can also be differentiated by multilocus sequence typing (MLST) (Gardan et al. 1999; Bull et al. 2011; Parkinson et al. 2011). Each genomospecies comprises strains, including the type strains that exhibit at least 70 % DNA-DNA homology. However, genomospecies can only be named formally when differentiating phenotypic characteristics are available (Wayne et al. 1987). Two strains were elevated to species status, *P. cannabina* and *P. tremae*, because phenotypic characteristics (e.g., unique carbon source utilization) were described that allowed these strains to be distinguished from other species, including *P. syringae* (Bull and Koike 2015; Marcelletti and Scortichini 2014). *Pseudomonas syringae* was first described as the causal agent of bacterial blight of onion in Japan (Goto 1972). Since then, bacterial blight of onion has been reported in many countries, including South Africa (Serfontein 2001), Japan (Kadota et al. 2000; Myung et al. 2011) and the United States of America (Sanders et al. 2003). Onion (*Allium cepa* L.) is an important vegetable crop worldwide and is grown by both commercial and emerging farmers. Onion is the third most important vegetable crop in South Africa, with regard to yield per hectare, total production and total value (SANSOR Annual Report, 2013). South Africa is one of the biggest producers of onion seed in the world. In the 2013/2014 growing season, 1514. 98 tons of onion seed were produced, of which 541 tons were for the export market (SANSOR Annual Report, 2015).

In 2007, bacterial blight of onion was reported in onion seed crops in Gauteng and Western Cape provinces of South Africa. The disease affected approximately 50 % of the crop. Symptoms in the field included irregular spots, each 4-10 cm long with a necrotic centre and water-soaked margin, on the leaves and seed stalks. Isolations were performed from the

symptomatic tissue. Isolations were also carried out from leek (*Allium porrum*) plants because bacterial blight symptoms also were observed in leek seed crops near the affected onion seed crops in the two provinces. Bacterial strains resembling *P. syringae* were isolated from both the onion and leek samples, and from onion seeds harvested from symptomatic seed stalks from the two provinces. The aim of this study was to characterize and confirm the identity of these onion and leek bacterial strains using Biolog GN III, MLST and pathogenicity tests.

2.3 Materials and methods

Bacterial strains

The reference strains used to compare utilization of carbon sources with Biolog GN III were obtained from the Collection Française de Bactéries associées aux Plantes (CFBP), France (Table 1). *Pseudomonas syringae* isolates from diseased onion plants and seed were obtained from the Agricultural Research Council (ARC) Plant Pathogenic and Plant Protecting Bacteria (PPPPB) National Collection in Pretoria, Gauteng, South Africa (Table 2). Three strains isolated from leek plants were included for comparison to 5 strains from onion plants and 18 strains from onion seed. Stock cultures of all strains were maintained in milk glycerol liquid medium at -80 °C. Preserved strains were transferred onto plates of King's B agar medium (containing the following per liter: proteose peptone No. 3 (Difco), 20 g; glycerol, 15 ml; K_2HPO_4 (anhydrous), 1.5 g; $MgSO_4 \times 7H_2O$, 1.5 g; and agar, 15 g) and incubated at 25 °C. Cultures checked routinely for purity and colony characteristics on King's B agar medium.

Biochemical and morphological tests

Gram strain reaction, production of fluorescent pigment on King's B agar medium, production of levan, oxidase activity, ability to cause potato rot, arginine utilization and production of a hypersensitive reaction on tobacco (LOPAT tests) were determined for each bacterial strain as described by Lelliott et al. (1966). Utilization of carbon sources such as erythritol, inositol, sorbitol, tyrosine and xylose was tested according to Young and Triggs (1994). The catabolic activity on 95 substrates of selected onion strains, including the pathotype and type strains of *P. syringae*, were determined using Biolog GN III Microplates (Biolog MicroLog Version 4.2 software; Biolog, Inc., Hayward, CA, USA), according to the manufacturer's instructions. Biolog data was entered into BioNumerics software Version 5.1 (Applied Maths, Kortrijk, Belgium). An unweighted pair group method with average linkage was constructed using the

Pearson's correlation coefficient.

Amplification and sequencing of *cts*, *gapA*, *gyrB* and *rpoD* genes

Genomic DNA of 13 onion strains and two references strains, *P. syringae* pv. *porri* CFBP 1908^{PT} and *P. syringae* CFBP 2336, was extracted with the Wizard Genomic DNA purification Kit (Promega, Madison, USA) according to the manufacturer's instructions. Purified DNA was quantified using a Dyna Quant 200 fluorometer (Nano drop, Hoefer, San Francisco, CA, USA). The DNA was stored at −20 °C until further analysis. MLST analysis was performed by sequencing four housekeeping genes: *cts* (citrate synthase), *gapA* (glyceraldehyde-3-phosphate dehydrogenase) *gyrB* (DNA gyrase B) and *rpoD* (RNA polymerase sigma70). The primers (Table S2a) used for PCR amplification and sequencing of the *cts*, *gapA* and *gyrB* genes were developed by Hwang et al. (2005), and the *rpoD* gene primers used were described by Sarkar and Guttman (2004). PCR amplification was performed as described previously by Morris et al. (2008). PCR products were separated on 1 % agarose gels at 80 V for 45 min, and purified using ExoSAP PCR cleanup reagent (Affymetrix, Danta Clara, CA, USA). PCR products were sequenced as described by Yan et al. (2008) by Inqaba Biotechnology (Pretoria, South Africa). Partial sequences from 13 *P. syringae* strains used for MLST analysis were submitted to the National Center of Biotechnology Information (NCBI) GenBank database, and accession numbers are available for the four DNA regions sequenced for each strain (Table 3).

Sequence analysis

The DNA sequences of the 13 *P. syringae* onion and leek strains were aligned using MAFFT (Version 7) online alignment tool (Katoh and Stanley 2013). Once aligned, the sequences were trimmed in BioEdit Sequence Alignment Editor (Hall 1999). A partition-homogeneity test was performed in PAUP 4.0b10 software (Swofford 2000) to establish if the four genes could be combined to form a single concatenated data set. The most suitable model was chosen for each dataset using the JModelTest Version 2.1.3 program (Posada 2008). Maximum Parsimony (MP) phylogenetic trees with 1000 replicates were constructed using PAUP 4.0b10 software. The trees were viewed and edited using MEGA version 5.05 (Tamura et al. 2011). Genetic relatedness of 13 of the 23 *P. syringae* strains associated with onion and leek bacterial blight was determined. The reference sequences (Table 1) were taken from the Plant Associated and Environmental Microbes Database (PAMDB, htt://genomeppws.vt.edu/cgi-

bin/MLST/home.pl; Almeida et al. 2010; Berge et al. 2014). The strain *P. graminis* 38Bb9 was used as the outgroup (Berge et al. 2014).

Pathogenicity tests

Pathogenicity screening of 23 onion and 3 leek strains used in the Biolog analysis was performed on seedlings of the onion cv. Granex 33, chive (*Allium schoenoprasum*) cv. Grasiue, leek (*Allium porrum* cv. Giant Italian and spring onion (*Allium fistulosum* cv. Salotte) in a glasshouse as described by Goszczynska et al. 2006. Briefly, a sterile needle was dipped into a bacterial colony of the appropriate strain growing on King's B agar medium (24-48 h of growth) and then inserted under the epidermis at one site on each of two leaves per plant. A total of 28 seedlings per plant species (= cultivar) were inoculated for each bacterial strain. Pathogenicity tests were repeated, with a total of 224 seedlings inoculated. Four plants per cultivar were inoculated with *P. syringae* pv. *porri* CFBP 1908^{PT} that served as the positive control treatment, and 4 plants of each species were used for the negative control treatment in which plants were injected similarly with sterilized water. Inoculated plants were incubated in a glasshouse with 27 °C/23 °C day/night temperature, and were observed daily for the development of symptoms. Spray inoculation was performed on eight onion and five leek strains listed in Table 4. The strains were grown on King's B agar medium at 28 °C for 48 h. Bacterial suspensions were made in sterilized distilled water (10^2, 10^5, 10^7 colony forming unit (CFU) per ml as determined by dilution plating. Six to eight week-old onion plants (*A. cepa* - cv. Granex 33) and leek plants (*A. porri* - cv. Giant Italian) were spray inoculated to runoff with each bacterial suspension. Negative control plants were inoculated with sterilized distilled water. Three plants were inoculated with each bacterial suspension. A total of 39 seedlings per cultivar were spray inoculated until runoff. Plants were incubated for 24 h in a humidity chamber at 27 °C and relative humidity of 95 %. Later, plants were maintained in a greenhouse with 27 °C/23 °C day/night temperature, and were observed daily for the development of symptoms. All inoculations were conducted twice. The symptoms were recorded as positive (+) corresponding to development of symptoms (water soaked spots, bleached or blight lesions), and negative (−) for no symptoms. Bacteria were re-isolated from the developing lesions on King's B agar. The identities of the bacteria re-isolated from lesions were confirmed by colony morphology on King's B agar, fluorescence under UV light, LOPAT tests and utilization of erythritol and sorbitol as single carbon sources. Re-isolation and results of the above tests fulfilled the Koch's postulates.

2.4 Results

Biochemical and morphological tests

All bacterial strains evaluated in this study (Table 2.2) were Gram negative rods, produced fluorescent pigment on King's B agar medium, tested positive for levan production, produced a hypersensitive reaction on tobacco, and tested negative for oxidase, arginine dihydrolase and potato rot (LOPAT tests). Carbon source utilization (Young and Triggs 1994) separated the *P. syringae* strains from onion into two groups (Table 2.2). Of the 23 strains tested, five did not utilize erythritol including three leek strains. Furthermore, Biolog GN III dendrograms (Fig. 2.1) also showed that the onion strains separated into two clusters. The first cluster contained 17 strains and did not group with any reference strains. The second cluster grouped four onion and two leek strains with *P. syringae* pv. *porri*, CFBP 2395 and CFBP 4235 (Fig. 2.1). In addition, Biolog results (Table S1) showed that the strains belonging to a new pathovar of P. syringae differed from the type strain of *P. syringae* CFBP 1392^T and *P. syringae* pv. *porri* strains by their ability to produce acid from erythritol and not utilize 3-methyl glucose, D-sorbitol and α-keto butyric acid (Table S1).

Multilocus sequence typing (MLST)

The sequences of the four genes *cts*, *gapA*, *gyrB* and *rpoD* were concatenated (1839 base pair). The concatenated tree (Fig. 2.2) supported the results of Biolog analysis (Fig. 2.1). The phylogenetic tree separated the onion strains into two groups. The strains in the first group formed a cluster without the reference strains, and the strains in the second group clustered with the reference strain *P. syringae* pv. *porri* CFBP 1908^{PT} and *P. syringae* CFBP 2336 with bootstrap support of 100 %. The reference strains of P. syringae clustered according to their phylogroups (Berge et al. 2014), with exception of *P. syringae* B64 which seem to be closely related to the onion and leek strains.

Pathogenicity tests

All 18 strains of a new pathovar tested for pathogenicity induced water soaked spots when inoculated directly onto onion seedlings, but did not cause symptoms on the leek seedlings. In contrast, the strains belonging to *P. syringae* pv. *porri* were pathogenic on both onion and leek seedlings. On leek seedlings, *P. syringae* pv. *porri* strains induced bleached to chlorotic

margins. No symptoms were observed on chive seedlings inoculated with strains of either pathogen. In the spray inoculation, eight strains of a new pathovar and five strains of *P. syringae* pv. *porri* were used. The strains of the new pathovar induced symptoms only on onion. Five strains of *P. syringae* pv. *porri* produced symptoms on both onion and leek. Symptoms induced of the new pathovar and five strains of *P. syringae* pv. *porri* in spray inoculation tests were identical not only on onion but also on leek. The water soaked and chlorotic lesions appeared at the tips of leaves and expanded into longitudinal blight (Fig. 2.3a-f). When the concentrations of 10^5 and 10^7 CFU/ml were used, the symptoms appeared five days after spraying. When the concentration of 10^2 CFU/ml was used, the symptoms started to develop 11 days after inoculation. Negative control plants sprayed with sterile distilled water did not develop any symptoms. To fulfil Koch's postulates, fluorescent *Pseudomonas syringae* colonies were re-isolated from all symptomatic plants on King's B medium. In LOPAT tests, all re-isolated strains were Levan positive and induced a hypersensitive reaction in tobacco. They were negative in oxidase, arginine dihydrolase and potato rotting tests. The re-isolated strains of the new pathovar utilised erythritol but not sorbitol. Negative control plants injected with sterilized distilled water did not develop symptoms. Fluorescent *P. syringae* colonies were re-isolated from symptomatic leaves of onion, leek and spring onion for both inoculation methods. The colonies were identified by LOPAT tests (Lelliott et al. 1966) and utilization of erythritol and sorbitol.

2.5 Discussion

Bacterial blight of onion was reported to be caused by *P. syringae* pv. *porri* and *X. axonopodis* pv. *allii* (Myung et al. 2011; Roumagnac et al. 2004; Serfontein 2001). In this study, the biochemical tests, Biolog GN III analysis and MLST based on *cts*, *gapA*, *gyrB* and *rpoD* genes successfully identified *P. syringae* strains from onion to pathovar levels. The Biolog dendrogram (Fig. 2.1) and MLST analysis (Fig. 2.2) revealed that bacterial blight of onion is caused by two pathovars: *P. syringae* pv. *allii*, the name proposed in this study, and *P. syringae* pv. *porri*. Both pathovars were isolated from onion plants and onion seed, and caused similar symptoms on inoculated onion plants to those observed in the field. Similarly, bacterial blight of pea can be caused by *P. syringae* pv. *pisi* or *P. syringae* pv. *syringae*, which produce indistinguishable symptoms in the field (Lawyer and Chun 2001).

Classification of *P. syringae sensu lato* into genomospecies and pathovars is still fraught with difficulty (Bull et al. 2011). Recently, however, MLST has been used to study the genetic

diversity and identify pathovars and strains within this complex (Berge et al. 2014; Bull et al. 2011; Cunty et al. 2015). MLST is convenient for identification of bacterial strains if the sequenced gene fragments are identical to those of the type and pathotype strains. Sarkar and Guttman (2004) were the first to provide MLST analysis of *P. syringae*, based on which they reported the strains of this species to comprise a highly clonal population. The authors used seven housekeeping genes to determine the evolutionary history of *P. syringae* strains covering the diversity of the species complex. Their results showed that the core genome of *P. syringae* is not susceptible to recombination or horizontal gene transfer. These authors further reported that *P. syringae* maintained a constant population size through time and that genetic variation in the core genome is associated weakly with the host plant species from which the strains were isolated. In this study, the same four genes (*cts*, *gapA*, *gyrB* and *rpoD*) evaluated by Hwang et al. (2005) were used to identify the pathovars of *P. syringae* infecting onion and leek plants in South Africa. Using four housekeeping genes ensured enough variability to differentiate closely related *Pseudomonas* strains (Bull et al. 2011; Ferrante and Scortichini 2014). Berge et al. (2014) provided "a user guide" to study the diversity and classification of *P. syringae* strains. The authors proposed using *cts* sequences as a rapid and precise means of classifying new strains. Citrase synthase is an enzyme present in almost all living cells and is central in a metabolic pathway, playing a key role in energy production and as a biosynthetic precursor (Wiegand and Remington 1986). The *cts* gene sequences in bacteria have minimum recombination events and the most congruence among the trees constructed with the four housekeeping genes evaluated in this study (Hwang et al. 2005). Pathogenicity test results on onion, leek and spring onion seedlings corresponded with those of Noble et al. (2006). Those authors, however, did not do pathogenicity tests on chive. The results of this study indicate that pathogenicity tests combined with molecular analysis were reliable for differentiating and classifying plant pathogenic bacteria isolated off symptomatic onion and leek plants and from onion seed. Stackebrandt et al. (2002) encouraged researchers to propose new species based upon other genomic methods, provided researchers can prove that, within the taxa studied, there is a sufficient degree of congruence among the techniques used. Based on Biolog GN III and MLST analysis in this study, some of the strains of *P. syringae* pv. *allii* obtained off onion plants and seed in South Africa are genetically and phenotypically different from other described pathovars of *P. syringae*. Furthermore, pathogenicity test results showed that the strains of *P. syringae* pv. *allii* induced symptoms only on onion, not on chive, leek, or spring onion. Thus, we conclude that *P. syringae pv. allii* is a new pathovar of *P. syringae sensu lato* that can cause bacterial leaf blight of onion.

Description of *Pseudomonas syringae* pv. *allii*

Pseudomonas syringae pv. *allii* adjective from *Allium*, the latin name of onion (*Allium cepa*). Colonies of *Pseudomonas syringae* pv. *allii* cultured on tryptone glucose extract agar plates are round, convex, viscous, translucent and white. Cells are Gram negative rods, aerobic, motile, produce fluorescent pigment on King's B medium. Positive reactions are: levan production, tobacco hypersensitive reaction, mannitol, inositol, sucrose, erythritol and xylose. Negative reactions are: oxidase, arginine hydrolase, potato rot test, sorbitol and tyrosine. *Pseudomonas syringae* pv. *allii* uses: sucrose, α-Dglucose, D-mannose, 1 % NaCI, 4 % NACI, 8 % NaCI, D-fructose, D-galactose, D-fucose, 1 % sodium lactate, fusidic acid, D-serine, D-mannitol, D-arabitol, glycerol, troleandomycin, rifamycin SV, minocycline, L-alanine, L-aspartic acid, L-glutamic acid, L-serine, lincomycin, niaproof 4, pectin, D-gluconic acid, D-glucoronic acid, glucoronamide, quinic acid, D-saccharides acid, vancomycin, tetrazolium violet, tetrazolium blue, α-ketoglutaric acid, nalidixic acid, Tween 40, γ-amino butyric acid, acetoacetic acid and aztreonam. The following carbon sources are not used by *Pseudomonas syringae* pv. *allii*: D-cellobiose, D-turanose, D-galactose, D-salicin, N-acetyl-β-D-mannosamine, N-acetyl-D-galactosamine, gelatin, D-aspartic acid, L-galactoni acid lactone and p-hydroxy-phenylacetic acid. The strains that formed a separate cluster from the reference strains and the strains that grouped with *Ps. syringae* pv. *porri* varied in utilizing dextrin, D-trehalose, gentiobiose, D-raffinose, D-malibiose, L-fucose, L-rhamnose, inosine, D-glucose-6-PO4, glycyl-Lproline, L-arginine, L-histidine, methyl pyruvate, L-lactice acid, β-hydroxy-D-L-butyric acid and acetoacetic acid. *Pseudomonas syringae* pv. *allii* utilizes erythritol, does not utilize sorbitol, methyl glucose, D-sorbitol and α-keto butyric acid. Pathogenic strains of *Pseudomonas syringae* pv. *allii* specifically induce water soaked lesion and leaf blight symptoms on onion plants. Three strains were deposited at Collection Française de Bactéries associées aux Plantes, France (CFBP). The pathotype strain of pathovar *allii* is BD 359 = 8400, and was deposited together with strains BD 346 = CFBP 8398 and BD 355 = CFBP 8399.

Acknowledgments

We thank the Agricultural Research Council: Plant Protection Research of South Africa for financial support, the Plant Phytobacteriology Group (PPG) at the Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, South Africa for assistance in editing the manuscript and Mr. Harry Boroko for technical assistance.

2.6 References

Almeida, N. F., Cai, R. S. Y., Clarke, R. C.,Morris, E. C., Schaad, W. N., Schuenzel, L. E., Lacy, H. G., Sun, X., Jones, B. J., Castillo, A. J., Bull, T. C., Leman, S., Guttman, S. D., Setubal, C. J., & Vinatzer, B. A. (2010). PAMDB, a multilocus sequence typing and analysis database and website for plant associated microbes. *Phytopathology*, 100, 208–215.

Berge, O., Monteil, L. C., Bartoli, C., Chandeysson, C., Guilbaud, C., Sands, C. D., & Morris, C. E. (2014). A user's guide to a data base of the diversity of Pseudomonas syringae and its application to classifying strains in this phylogenetic complex. *PloS One*, 9(9), e105547.

Bull, C. T., & Koike, S. T. (2015). Practical benefits of knowing the enemy: Modern molecular tools for diagnosing the etiology of bacterial diseases and understanding the taxonomy and diversity of plant-pathogenic bacteria. *Annual Review of Phytopathology*, 53, 157–180.

Bull, C. T., Clarke, C. R., Cai, R., Vinatzer, B. A., Jardini, T. M., & Koike, S. T. (2011). Multilocus sequence typing of *Pseudomonas syringae sensu lato* confirms previously described genomospecies and permits rapid identification of *P. syringae* pv. *coriandricola* and *P. syringae* pv. *apii* causing bacterial leaf spot on parsley. *Phytopathology*, 101, 847–858.

Cunty, A., Poliakoff, F., Rivoal, C., Cesbron, S., Fischer-Le Saux, M., Lemaire, C., Jacques, A. M., Mancaue, C., & Vanneste, J. L. (2015). Characterization of *Pseudomonas syringae pv. actinidiae* (Psa) isolated from France and assignment of Psa biovar 4 to a de novo pathovar: *Pseudomonas syringae* pv. *actinidifoliorum* pv. nov. *Plant Pathology,* 64, 582–596.

Dye, D. W., Bradbury, J. F., Goto, M., Hayward, R. A., Lelliott, R. A., & Schroth, M. N. (1980). International standards for naming pathovars of phytopathogenic bacteria and a list of pathovar names and pathotype strains. *Revision of Plant Pathology*, 59, 153–168.

Ferrante, P., & Scortichini, M. (2014). Redefining the global populations of *Pseudomonas syringae* pv. *actinidiae* based on pathogenic, molecular and phenotypic characteristics. *Plant Pathology*, 64, 51–62.

Gardan, L., Shafik, H., Belouin, S., Broch, R., Grimont, F., & Grimont, P. A. D. (1999). DNA relatedness among the pathovars of *Pseudomonas syringae* and description of *Pseudomonas tremae* sp. nov. and *Pseudomonas cannabina* sp. nov. (exSutic and Dowson 1959). *International Journal of Systematic Bacteriology*, 49, 469–478.

Goszczynska, T., Moloto, V. M., Venter, S. N., & Coutinho, T. A. (2006). Isolation and identification of *Pantoea ananatis* from onion seed in South Africa. *Seed Science and Technology*, 34, 677–690.

Goto, M. (1972). Bacterial leaf blight of onion in Japan. Plant Disease Report, 56, 490–493.

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for window 95/98/NT. *Nucleic Acids Symposium Series*, 41, 95–98.

Hwang, M. S. H., Morgan, R. L., Sarkar, S. F., Wang, P. W., & Guttman, D. S. (2005). Phylogenetic characterization of virulence and resistance phenotypes of *Pseudomonas syringae*. *Applied and Environmental Microbiology*, 71, 5182–5191.

Kadota, I., Uehara, K., Shinohara, H., & Nishiyama, K. (2000). Bacterial blight of Welsh onion: A new disease caused by *Xanthomonas campestris* pv. *allii* pv. nov. *Journal of General Plant Pathology*, 66, 310–315.

Katoh, K., & Stanley, D. M. (2013). MAFFT Multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30, 772–780.

Lamichhane, J. R., Messëan, A., & Morris, C. E. (2015). Insights into epidemiology and control of diseases of annual plants caused by the *Pseudomonas syringae* species complex. *Journal of General Plant Pathology*, 81, 331–350.

Lawyer, A.S., & Chun, W. (2001). Foliar diseases caused by bacteria. (In Compendium of Pea Diseases, 2nd Ed. J.M. Kraft and F.L. Pfleger, Eds. *American Phytopathological Society* (pp. 22–24). St. Paul, MN).

Lelliott, R. A., Billing, E., & Hayward, A. C. (1966). A determinative scheme for the fluorescent plant pathogenic pseudomonads. *Journal of Applied Bacteriology*, 29, 470–489.

Marcelletti, S., & Scortichini, M. (2014). Definition of plant pathogenic *Pseudomonas* genomospecies of the *Pseudomonas syringae* complex through multiple comparative approaches. *Phytopathology*, 104, 1274–1282.

Morris, C. E., Sands, D. C., Vinatzer, B. A., Glaux, C., Guilbaud, C., Buffiere, A., Yan, S., Domiguez, H., & Thompson, M. B. (2008). The life history of the plant pathogen *Pseudomonas syringae* is linked to the water cycle. *International Society for Microbial Ecology*, 2, 321–334.

Myung, I. S., Joa, J. H., & Shim, H. S. (2011). Bacterial leaf spot of onion caused by *Pseudomonas syringae* pv. *porri,* a new disease in Korea. *Plant Disease*, 95, 1311.

Noble, H. D., Cother, J. E., Hailstones, L. D., Flack, M., Oxspring, L., & Hall, B. (2006).

Characterisation of *Pseudomonas syringae* strains associated with a leaf disease of leek in Australia. *European Journal of Plant Pathology*, 115, 419–430.

Palleroni N. J. (1984). Pseudomonads. (In: Krieg NR, ed. *Bergey's Manual of Systematic Bacteriology* (pp. 141–199). Williams and Wilkins, Baltimore).

Parkinson, N., Bryant, R., Bew, J., & Elphinstone, J. (2011). Rapid phylogenetic identification of members of the *Pseudomonas syringae* species complex using the *rpoD* locus. *Plant Pathology*, 60, 338–344.

Posada, D. (2008). JModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, 25, 1253–1256.

Roumagnac, P., Gagnevin, L., Gardan, L., Sutra, L., Manceau, C., Dickstein, R. E., Jones, B. J., Rott, P. & Pruvost, O. (2004). Polyphasic characterization of xanthomonads isolated from onion, garlic and Welsh onion (*Allium* spp.) and their relatedness to different *Xanthomonas* species. *International Journal of Systematic and Evolutionary Microbiology*, 54, 15–24.

Sanders, F. H., Langston Jr., B. D., Brock, H. J., Gitaitis, D. R., Curry, E. D., & Torrance, L. R. (2003). First report of a leaf blight of onion caused by *Xanthomonas* spp. in Georgia. *Plant Disease*, 87, 749.

SANSOR Annual Report 2011/2012 (2013). South African National Seed Organisation. Pretoria: South Africa.

SANSOR Annual Report 2013/2014 (2015). South African National Seed Organisation. Pretoria: South Africa.

Sarkar, S. F., & Guttman, D. S. (2004). Evolution of the core genome of *Pseudomonas syringae*, a highly clonal, endemic plant pathogen. *Applied and Environmental Microbiology*, 70, 1999–2012.

Serfontein, J. J. (2001). *Xanthomonas* blight of onion in South Africa. Plant Disease, 85, 442.

Stackebrandt, E., Frederiksen, W., Garrity, G. M., Grimont, P. A. D., Kämpfer, P., Maiden, M. C. J., Nesme, X., Rosselló- Mora, R., Swings, J., Trüper, H. G., Vauterin, L., Ward, A. C., & Whitman, W. B. (2002). Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. *International Journal of Systematic and Evolutionary Microbiology*, 52, 1043–1047.

Swofford, D. L. (2000). PAUP*: Phylogenetic Analysis Using Parsimony and other methods (software). Sutherland, MA: Sinauer Associates.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Sudhir Kumar, S. (2011) MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood,

Evolutionary Distance and Maximum Parsimony Methods. *Molecular Biology and Evolution,* 28, 2731–2739.

Van Hall, C. J. J. (1902). Bijdrage tot de kennis der bacteriëleplantenziekten. Companyöperatie vedrukkerijvereeniging “Plantijn”. Inaug Diss, Amsterdam 8, 372 (in Dutch).

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

Wiegand, G., & S. J. Remington, J. S. (1986). Citrate synthase: structure, control, and mechanism. *Annual Review of Biophysics and Biophysical Chemistry*, 15, 97–117.

Yan, S., Liu, H., Mohr, T. J., Jenrette, J., Chiodini, R., Zaccardelli, M., Setubal, J. C., & Vinatzer, B. A. (2008). Role of recombination in the evolution of the model plant pathogen *Pseudomonas syringae* pv. *tomato* DC3000, a very atypical tomato strain. *Applied and Environmental Microbiology*, 74, 3171–3181.

Young, J. M. (2010). Taxonomy of *Pseudomonas syringae*. *Journal of Plant Pathology*, 92, S1.5–S1.14.

Young, J.M., & Triggs, C.M. (1994). Evaluation of determinative tests for pathovars of *Pseudomonas syringae* van Hall 1902. *Journal of Applied Bacteriology*, 77, 195–207.

2.7 Tables

Table 2.1 Lists of references strains of *Pseudomonas* used for Biolog GN III and multilocus sequence typing analysis (MLST). The sequences for MLST were selected from Berge et al. (2014). All the sequences were obtained from the Plant associated and environmental microbes database (Almeida et al., 2010).

Strain number	Strain name	Host plant or medium	Phylogroup[c]
1_6	*P. syringae* pv. *oryzae*	*Oryza sativa*	4
38B9	*P. graminis*	Water phase of a cloud	nd
AI0084	*P. syringae*	Stream water	2
AI0095	*P. syringae*	Stream water	10
AI0105	*P. syringae*	Stream water	10
B64	*P. syringae* pv. *syringae*	*Triticum aestivum*	2
B728a	*P. syringae* pv. *syringae*	*Phaseolus vulgaris*	2
CC1430	*P. syringae*	Epilithic biofilm	1
CC1431	*P. syringae*	Epilithic biofilm	1
CC1457	*P. syringae*	*Primula* sp.	2
CC1470	*P. syringae*	Stream water	2
CC1582	*P. syringae*	Epilithic biofilm	8
CCV0055	*P. syringae*	Stream water	13
CFBP 1392[a,b]	*P. syringae* type strain	*Syringa vulgaris*	2
CFBP 1670[a]	*P. savastanoi* pv. *savastanoi*	*Olea europaea*	3
CFBP 1908[a,b]	*P. syringae* pv. *porri*	*Allium porrum*	4
CFBP 2104[a]	*P. syringae* pv. *lachrymans*	*Cucumis sativus*	1
CFBP 2212[a]	*P. syringae* pv. tomato	*Solanum lycopersicum*	1
CFBP 2216[a]	*P. syringae* pv. *coranafaciens*	*Avena sativa*	4
CFBP 2336[a,b]	*P. syringae*	*Allium cepa*	nd[d]

CFBP 2337[a]	*P. syringae*	*A. cepa*	nd
CFBP 2353[a]	*P. syringae* pv. *theae*	*Thea sinensis*	1
CFBP 2395[a]	*P. syringae* pv. *porri*	*A. porrum*	4
CFBP 4235[a]	*P. syringae* pv. *porri*	*A. porrum*	4
CFBP 4702[a]	*P. syringae* pv. *syringae*	*S. vulgaris*	2
CFBP 2067[a,b]	*P. syringae* pv. *helianthi*	*Helianthus annuus*	6
CFBP 4407	*P. cichorii*	*Lactuca sativa*	11
CMO0010	*P. syringae*	Rain	3
CMO0085	*P. syringae*	Rain	9
CST0099	*P. syringae*	Rain	9
CSZ0260	*P. syringae*	Stream water	7
DC3000	*P. syringae* pv. *tomato*	*S. lycopersicum*	1
GAW0113	*P. syringae*	Irrigation canal water	12
KN203	*P. syringae* pv. *maculicola*	*Brassica rapa*	1
KN221	*P. syringae* pv. *coronafaciens*	*A. sativa*	4
LAB0163	*P. syringae*	Epilithic biofilm	8
MAFF301020	*P. syringae* pv. *mori*	*Morus alba*	3
MAFF301765	*P. syringae* pv. *glycine*	*Glycine max*	3
PsyCit7	*P. syringae*	*Citrus sinensis*	2
SZ0131	*P. syringae*	Stream water	13
USA0046	*P. syringae*	Stream water	7

[a] Bacterial strains tested with Biolog GN III only.

[b] Bacterial strains tested with both Biolog GN III and the four housekeeping genes: citrase synthase (*cts*), glyceraldehyde-3-phosphate dehydrogenase (*gapA*), DNA gyrase B (*gyrB*) and RNA polymerase sigma^{70} (*rpoD*).

[c] Group of organisms determined based on multilocus sequence typing (MLST) phylogenetic tree (Berge et al. 2014). nd - not determined.

Table 2.2 Phenotypic tests that differentiate strains of *P. syringae* pv. *allii*, *P. syringae* pv. *porri* and *P. syringae* pv. *syringae*. Carbon sources utilization as described by (Young & Triggs, 1994).

Bacterial strain[a]	Original host plant	Differential carbon sources				
		Inositol	Erythritol	Sorbitol	Tyrosine	Xylose
P. syringae* pv. *allii						
BD 341	Onion seed	+	+	-	-	+
BD 345	Onion seed	+	+	-	-	+
BD 346	Onion seed	+	+	-	-	+
BD 347	Onion seed	+	+	-	-	+
BD 349	Onion seed	+	+	-	-	+
BD 350	Onion seed	+	+	-	-	+
BD 351	Onion seed	+	+	-	-	+
BD 353	Onion seed	+	+	-	-	+
BD 354	Onion seed	+	+	-	-	+
BD 355	Onion seed	+	+	-	-	+
BD 356	Onion seed	+	+	-	-	+
BD 357	Onion seed	+	+	-	-	+
BD 358	Onion seed	+	+	-	-	+
BD 359	Onion seed	+	+	-	-	+
BD 360	Onion seed	+	+	-		+
BD 406	Onion plant	+	+	-	-	+
BD 407	Onion plant	+	+	-	-	+
BD 410	Onion plant		+	-	-	+
P. syringae* pv. *porri		+				
BD 370	Leek	+	-	+	+/-	+
BD 371	Leek	+	-	+	+/-	+
BD 374	Leek	+	-	+	+/-	+
BD 424	Onion stalk	+	-	+	+/-	+
BD 431	Onion stalk	+	-	+	+/-	+
BD 445	Onion seed	+	-	+	+/-	+
BD 447	Onion seed	+	-	+	+/-	+
BD 450	Onion seed	+	-	+	+/-	+
CFBP 1908[PT]	Leek		-	+	+/-	+

P. syringae* pv. *syringae		+				
CFBP 1392^T	Lilac		+	+	-	-

[a] BD strains: accession numbers of the Plant Pathogenic and Plant Protecting Bacteria (PPPPB) culture collection, Agricultural Research Council – Plant Protection Research Institute, Pretoria, South Africa. CFBP strain obtained from the Collection Française de Bactéries associées aux Plantes, France.

+, the carbon source is utilized

+/-, the carbon source is weakly utilized.

-, the carbon source is not utilized.

Table 2.3. GenBank Accession Numbers of the DNA sequences of four housekeeping genes amplified from *Pseudomonas* strains and used for phylogenetic analysis of bacteria isolated from onion and leek seed crops in South Africa that had symptoms of bacterial leaf blight, and from onion plant and seed produced in South Africa.

Strain	**Original host plant**	***cts* partial sequence**	***gapA* partial sequence**	***gyrB* partial sequence**	***rpoD* partial sequence**
BD 341	Onion seed	KT288088	KT288097	KT288079	KT288070
BD 345	Onion seed	KT288089	KT288098	KT288080	KT288071
BD 346	Onion seed	KT288090	KT288099	KT288081	KT288072
BD 351	Onion seed	KT288091	KT288100	KT288082	KT288073
BD 353	Onion seed	KT288092	nd	KT288083	KT288074
BD 355	Onion seed	KT288093	KT288101	KT288084	KT288075
BD 359	Onion seed	KT288094	KT288102	KT288085	KT288076
BD 370	Leek	KT328506	KT328500	KP698142	KM873343
BD 374	Leek	KT799564	KT799569	KP698146	KM873347
BD 410	Onion plant	KT288095	KT288103	KT288086	KT288077
BD 424	Onion plant	KT328507	KT328501	KP698148	KM873349
BD 431	Onion plant	KT328508	KT328502	KP698153	KM873355
BD 447	Onion plant	KT328509	KT328503	KT799576	KT799580
CFBP 2336	Onion	KT799566	KT799572	nd	KM873361
CFBP 1908^{PT}	Leek	KT799565	KT799571	nd	nd

nd not done

Table 2.4. Results of pathogenicity tests observed after 3 days to 11 days after inoculating *Pseudomonas syringae* strains onto seedlings of onion (*Allium cepa*), chive (*Allium schoenoprasum*), leek (*Allium porrum*), and spring onion (*Allium fistulosum*).

Bacterial strain[a]	**Original host plant**	**Stab inoculation test**				**Spray inoculation test**	
		Onion cv. Granex 33	**Chive cv. Grasiue**	**Leek cv. Italian Giant**	**Spring onion cv. Salotte**	**Onion cv Granex 33**	**Leek cv. Italian Giant**
P. syringae* pv. *allii							
BD 341	Onion seed	+	-	-	-	+	-
BD 345	Onion seed	+	-	-	-	+	-
BD 346	Onion seed	+	-	-	-	+	-
BD 351	Onion seed	+	-	-	-	+	-
BD 353	Onion seed	+	-	-	-	+	-
BD 354	Onion seed	+	-	-	-	+	-
BD 355	Onion seed	+	-	-	-	+	-
BD 359	Onion seed	+	-	-	-	+	-
BD 406	Onion plant	+	-	-	-	nd	nd
BD 407	Onion plant	+	-	-	-	nd	nd
BD 410	Onion plant	+	-	-	-	+	-
BD 347	Onion seed	+	-	-	-	nd	nd
BD 349	Onion Seed	+	-	-	-	nd	nd
BD 350	Onion seed	+	-	-	-	nd	nd
BD 356	Onion seed	+	-	-	-	nd	nd

BD 357	Onion seed	+	-	-	-	Nd	nd
BD 358	Onion seed	+	-	-	-	Nd	nd
BD 360	Onion seed	+	-	-	-	Nd	nd
P. syringae* pv. *porri							
BD 370	Leek	+	-	+	+	+	+
BD 374	Leek	+	-	+	+	+	+
BD 424	Onion stalk	+	-	+	+	+	-
BD 431	Onion stalk	+	-	+	+	+	-
BD 447	Onion stalk	+	-	+	+	+	+
BD 445	Onion stalk	+	-	+	+	Nd	Nd
BD 450	Onion stalk	+	-	+	+	Nd	Nd
BD 371	Leek	+	-	+	+	Nd	Nd
CFBP 1908[PT]	Leek	+	-	+	+	+	+
Negative control	Water	-	-	-	-	-	-

[a] BD strains: accession numbers of the Plant Pathogenic and Plant Protecting Bacteria (PPPPB) culture collection, Agricultural Research Council – Plant Protection Research Institute, Pretoria, South Africa. CFBP strain obtained from the Collection Française de Bactéries associées aux Plantes, France. Pathogenicity tests on 8 onions, 5 leeks, a positive control and negative control were done in the glasshouse.

[b] the seedlings were spray inoculated with different concentrations (10^2, 10^5 and 10^7 CFU/ml as determined by dilution plating). Similar results were obtained on onion and leek.

[c] the seedlings were inoculated by injecting the epidermis of the leaves with a sterile needle dipped into a bacterial colony. The strains of a new pathovar were pathogenic on onion only. The strains of *P. syringae* pv. *porri* were pathogenic on onion, leek and spring onion. No symptoms were observed on chives.

+, strain produced water soaked or leaf blight lesion on the seeding. -, strain did not produce symptom on the seedling.

2.8 Figures

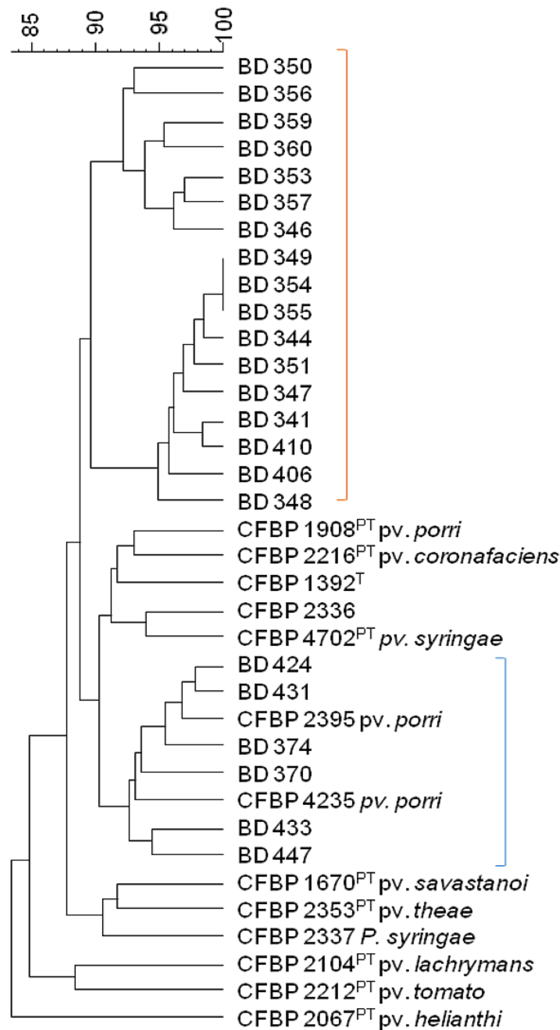


Figure 2.1 Relationships among 36 *Pseudomonas syringae* isolated from onion, including reference strains of *P. syringae*, based on Biolog GN III Microplate substrate utilization patterns. *Pseudomonas syringae* isolated from onion and leek plants are shown in red and blue, respectively. Pearson's correlation coefficient was used to construct an unweighted paired grouped method with arithmetic mean dendrograms using BioNumerics software (version 4.5, Applied Maths, Kortrijk, Belgium).

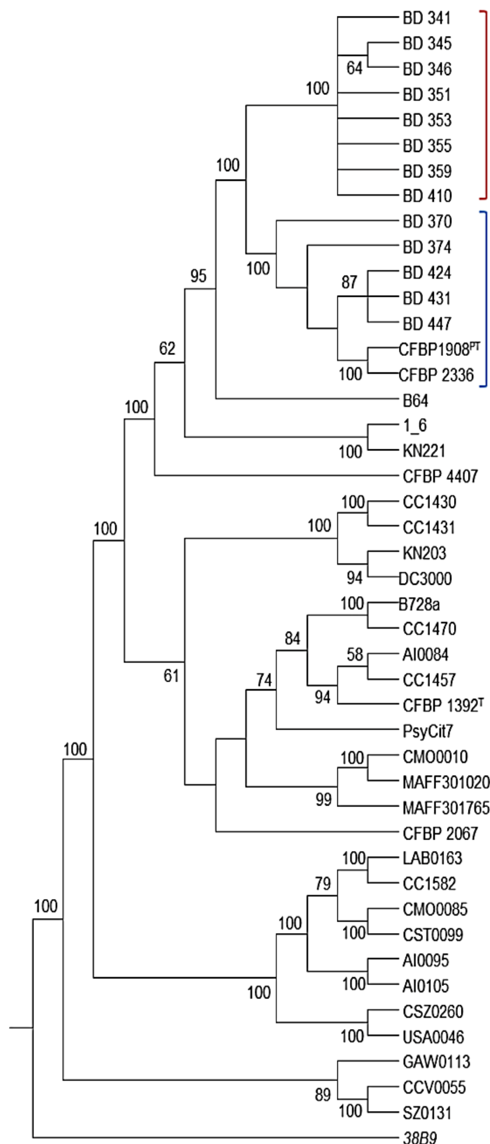


Figure 2.2 Maximum parsimony phylogenetic tree of 45 *Pseudomonas syringae* strains based on partial DNA sequences of four housekeeping genes: a) *cts*, b) *gapA*, c) *gyrb*, and d) *rpoD*. Bootstrap values of 1000 replicates were applied. Eight strains shown in red belong to *P. syringae* pv. *allii* and five strains shown in blue belong to *P. syringae* pv. *porri*. Sequences of the references strains were obtained from the Plant associated and environmental microbes database (Berge et al. 2014).

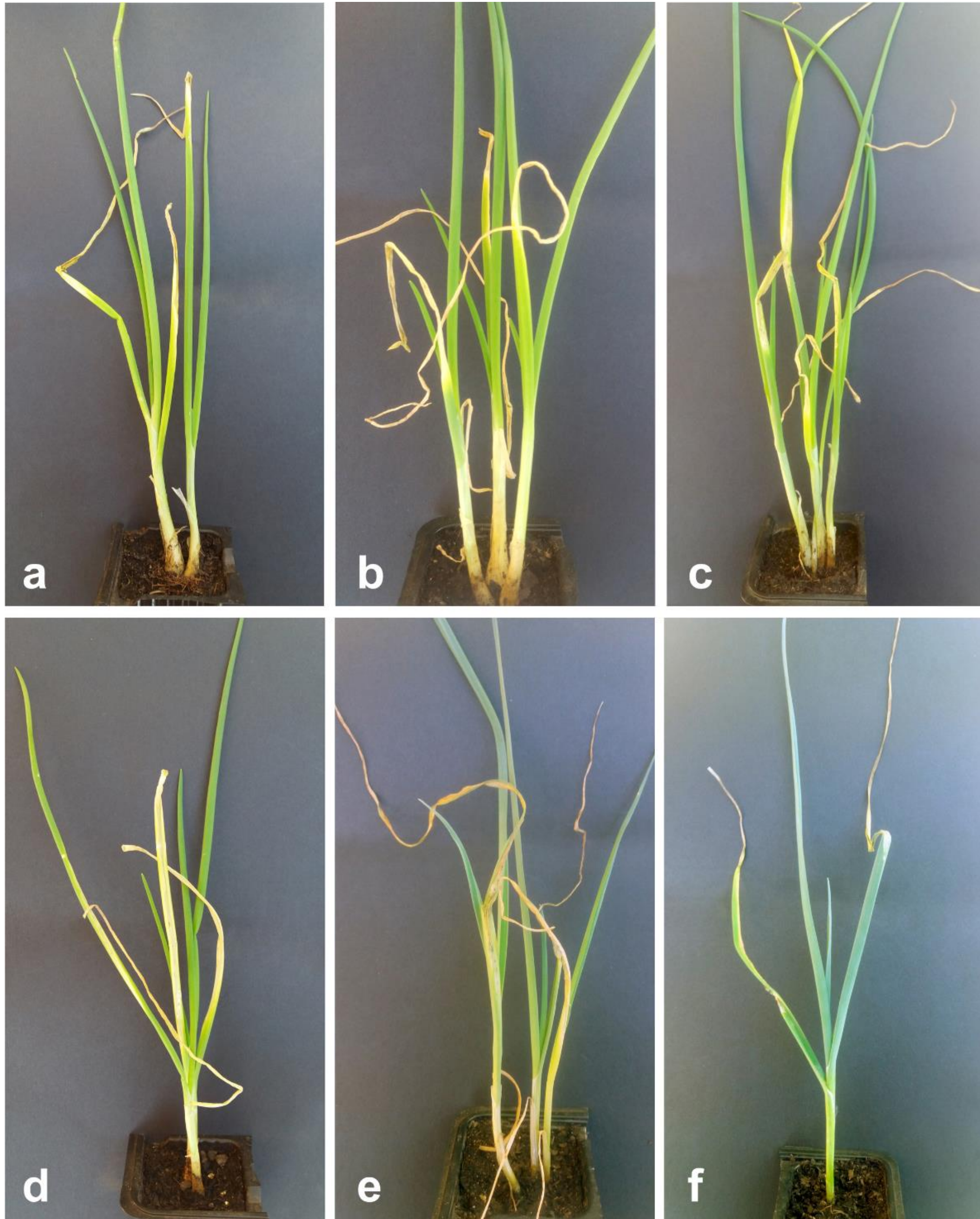


Figure 2.3 a) Water-soaked lesion induced by *Pseudomonas syringae* pv. *allii* on onion (*Allium cepa*), b) necrotic lesion induced by *P. syringae* pv. *porri* on a leek (*Allium porrum*) seedling, c) necrotic lesion induced by *P. syringae* pv. *porri* on onion, and d-f) necrotic lesions on spring onion leaves induced by *P. syringae* pv. *porri*.

2.9 Online supplementary materials

Table S1a. Utilization of 95 carbon sources by four strains of *Pseudomonas syringae* pv. *allii*, the type strain of *P. syringae* CFBP 1392^{T}, *P. syringae* pv. *porri* CFBP 1908^{PT} and strain CFBP 2337 isolated from onion plants, tested with Biolog GN III Microplates. Carbon sources highlighted in grey distinguish strains of *P. syringae* pv. *allii* from strains CFBP 1392^{T}, CFBP 1908^{PT} and CFBP 2337.

Carbon sources	**CFBP 1392^{T}**	**CFBP 1908^{PT}**	**CFBP 2337**	**BD 346**	**BD 355**	**BD 359**	**BD 410**
Dextrin	+/-	+	+	+	+	+	+
D-Maltose	-	-	-	-	-	-	-
D-Trehalose	-	-	-	-	-	-	-
D-Cellobiose	-	-	-	-	-	-	-
Gentiobiose	-	-	+	-	+	+	+
Sucrose	+	+	+	+	+	+	+
Turanose	-	-	-	-	-	-	-
Stachyose	-	-	-	-	-	-	-
pH 8	+	+	+	+	+	+	+
pH 5	+	+	+	+	+	+	+
D-Raffinose	-	-	-	-	-	-	-
α-D-Lactose	-	+	-	-	-	-	-
D-Melibiose	-	+	+	+	+	+	+
β-Methyl-D-Glucoside	-	-	-	-	-	-	-
D-salicin	-	-	-	-	-	-	-
N-Acetyl-D-Galactosamine	-	-	-	-	-	-	-
N-Acetyl-β-D-Mannosamine	-	-	-	-	-	-	-
N-Acetyl D-Galactosamine	-	-	-	-	-	-	-
N-Acetyl-Neuraminic acid	-	-	-	-	-	-	-
1 % NaCl	+	+	+	+	+	+	+
4 % NaCl	+	+	+	+	+	+	+
8 % NaCl	+	+	+	+	+	+	+
α-D-Glucose	+	+	+	+	+	+	+
D-Mannose	+	+	+	+	+	+	+
D-Fructose	+	+	+	+	+	+	+
D-Galactose	+	+	+	+	+	+	+

3-Methyl Glucose	+	+	+	-	-	-	-
D-Fucose	+	+	+	+	+	+	+
L-Fucose	+	+	+	+	+	+	+
L-Rhamnose	+	+	+	-	+	-	-
Inosine	+	+	+	+	-	-	-
1 % Sodium Lactate	+	+	+	+	+	+	+
Fusidic Acid	+	+	+	+	+	+	+
D-serine	+	+	+	+	+	+	+
D-Sorbitol	+	+	-	-	-	-	-
D-Mannitol	+	+	+	+	+	+	+
D-Arabitol	+	+	+	+	+	+	+
Myo-Inositol	+	+	+	-	-	+	-
Glycerol	+	+	+	+	+	+	+
D-Glucose-6-PO4	+	+	+	-	+	-	-
D-Fructose-6-PO4	-	+	+	+	+	+	-
D-Aspartic Acid	+	-	-	-	-	-	-
Troleandomycin	+	+	+	+	+	+	+
Rifamycin SG	+	+	+	+	+	+	+
Minocycline	+	+	+	+	+	+	+
Gelatin	-	-	-	-	-	-	-
Glycyl-L-Proline	+	-	+	-	+/-	+	+/-
L-Alanine	+	+	+	+	+	+	+
L-Arginine	+	-	-	+	+/-	+	+/-
L-Aspartic Acid	+	+	+	+	+	+	+
L-Glutamic Acid	+	+	+	+	+	+	+
L-Histidine	-	-	-	-	-	+	+
L-Pyroglutamic Acid	-	-	-	-	-	-	-
L-serine	+	+	+	+	+	+	+
Lincomycin	+	+	+	+	+	+	+
Guanidine HCl	-	+	+	+	+	+	+
Niaproof 4	+	+	+	+	+	+	+
Pectin	+	+	+	+	+	+	+
D-Galactonic Acid	+	+	+	+	+	+	+
L-Galactonic Acid Lactone	-	-	-	-	-	-	-
D-Gluconic Acid	+	+	+/-	+	+	+	+
D-Glucoronic Acid	+	+	+	+	+	+	+
Glucoronamide	+	+	+	+	+	+	+
Mucic Acid	+	+	+	+	+	+	+
Quinic Acid	+	+	+	+	+	+	+

D-Saccharic Acid	+	+	+	+	+	+	+
Vancomycin	+	+	+	+	+	+	+
Tetrazolium Violet	+	+	+	+	+	+	+
Tetrazolium Blue	+	+	+	+	+	+	+
P-Hydroxy-phenylacetic Acid	-	-	-	-	-	-	-
Methyl Puruvate	+	+	+	+	+	+	-
D-Lactic Acid Methyl Ester	-	-	-	-	-	-	-
L- Lactic Acid	-	-	-	-	-	-	-
Citric Acid	+	+	+	+	+	+	+
α-Keto-Glutaric Acid	+	+	+	+	+	+	+
D-Malic Acid	+	+	+	+	+	+	+
L-Malic Acid	+	+	+	+	+	+	+
Bromo-Succinic Acid	+	+	+	+	+	+	+
Nalidixic Acid	+	+	+	+	+	+	+
Lithium Chloride	+	+	+	+	+	+	+
Potassium Tellurite	+	+	+	+	+	+	+
Tween 40	+	+	+	+	+	+	+
ϒ-Amino Butyric Acid	+	+	+	+	+	+	+
α-Hydroxy Butyric Acid	+	-	-	-	-	-	-
β-Hydroxy-D-L-Butyric Acid	+	-	-	-	-	-	-
α-Keto-Butyric Acid	+	+	+	-	-	-	-
Acetoacetic Acid	+	+	+	+	-	+	+
Propionic Acid	+	+	+	-	+	+	+
Acetic Acid	+	+	+	+	+	-	+
Formic Acid	+	+	-	-	+	-	+
Aztreonam	+	+	+	+	+	+	+
Sodium Butyrate	+	+	+	+	+	+	+
Sodium bromate	+	-	-	-	-	-	-

+, carbon source was used.

+/-, carban source was weekly used.

-, carbon source was not used.

Table S1b. Multilocus sequencing type (MLST) primers used in this study to amplify four housekeeping genes from bacterial strains isolated from onion and leek seed crops with symptoms of bacterial blight, and from onion seed in South Africa.

Primers	Sequence	Reference
cts+174p	5'-GCCTCBTGCGAGTCGAAGATCACC -3'	Hwang et al. 2005
cts-1130s	5'-CGAAGATCACGG TGAACATGCTGG-3'	
gapA+264p	5'-CCGGCSGARCTGCCSTGG-3'	Hwang et al. 2005
gapA-312s	5'-TCGARTGCACSGGBCTSTTCACC-3'	
gyrB+271p	5'-TCBGCRGCVGARGTSATCATGAC-3'	Hwang et al. 2005
gyrB-1027s	5'-TTGTCYTTGGTCTGSGAGCTGAA-3'	
rpoDFp	5'-AAGGCGARATCGAAATCGCCAAGCG -3'	Sarkar & Guttman 2004
rpoDRps	5'-GGAACWKGCGCAGGAAGTCGGCACG -3'	

Chapter 3

Community analysis of bacteria associated with onion (*Allium cepa* L.) seed lots produced in South Africa

3.1 Abstract

Seeds can be involved in the transmission of microorganisms from one generation to another and, consequently, can act as reservoirs of plant microbiota. However, little is known about the structure of seed-associated bacterial assemblages and the regulators of the bacterial assemblage structure. In this study, the bacterial diversity on and within onion seed lots was examined using culture-dependent and culture-independent methods. Eighteen seed lots representing a single cultivar were obtained from the Northern and Western Cape Provinces, South Africa, the regions where the seed were produced. For the culture-dependent method, culturing was performed on tryptone glucose extract agar medium. Selected colonies were identified by sequence analyses of the 16S rRNA gene. Culturable isolates belonged to the phylum Proteobacteria and included representatives of *Acinetobacteria*, *Enterobacter*, *Erwinia*, *Microbacterium*, *Pantoea* and *Pseudomonas*. The bacteria of the genus *Pantoea* were isolated from all 18 seed lots. Total DNA extraction from the seed lots followed by 16S rRNA gene profiling was used as the culture-independent method. Results revealed that the Proteobacteria were the dominant phylum, with an average of 96.2% for all 18 samples. Bacteriocedetes, Firmicutes, Plantomycetes, Actinobacteria and Acinetobacteria were also detected, but in limited abundances from the 18 seed lots. The Greengenes taxonomic database detected 512 operational taxonomic units (OTUs) belonging to 6 phyla, 11 families and 7 genera. In general, both methods used to study bacterial diversity associated with these onion seed lots produced in South Africa revealed that *Pantoea* was the most prevalent genus detected from all 18 seed lots. This is the first study to determine bacterial communities associated with onion seeds in South Africa.

3.2 Introduction

Onions are regarded as the the third most important vegetable crop, in South Africa (Department of Agriculture, Forestry and Fisheries, 2015; The National Agricultural Directory, 2011). They are grown commercially in South Africa, mainly in the Western Cape, Northern Cape, Free State, North West and Limpopo Provinces (Department of Agriculture, Forestry and Fisheries, 2015). Onions are an important component of countless cuisines, lending varied flavours to meals, both cooked and raw (Benkeblia & Lanzotti, 2007; http:www.onions-usa.org). Research has shown that, onions play a significant role in preventing and treating human diseases (Van der Meer, 1997).

In the 2014/2015 growing season, a total volume of 724.80 tons of onions were produced, of which 657.07 tons were for the export market (SANSOR Annual Report, 2015). Seeds serve as an important means of agricultural production for many plant species. Studies have confirmed that the surface and interior of seeds harbour a variety of microbes, including beneficial and plant pathogenic bacteria as well as human pathogens (Adam et al., 2016; Barret et al., 2015; Guan, 2009; Hu et al., 2004; Johnston-Monje et al., 2016; National Advisory Committee on Microbiology Criteria for Food, 1999; Nelson, 2004). However, little is known about the structure of seed-associated microbial communities (Barret et al., 2015; Lopez-Velasco et al., 2013; Klaedtke et al., 2015; Rezki et al., 2016). Research has shown that microorganisms use three pathways to colonise seeds (Maude, 1996). For example, the floral pathway contribute to the transmission of plant growth promoting bacteria, plant pathogens and endophytes (Darrasse et al., 2010; Johnston-Monje & Raizada, 2011; Spinelli et al., 2005; Terrasson et al., 2015). Dutta et al. (2014) reported that seedborne bacterial pathogens can develop either on or in the seed coat or within the embryo as a result of this transmission pathway.

The movement of infected seed and subsequent seed transmission of plant pathogenic microorganisms represents an important means of dispersion of seedborne plant pathogens (Baker & Smith, 1966; Barret et al., 2016). Managing seedborne bacterial pathogens are difficult as the control measures are limited and more often ineffective (Darrasse et al., 2010; Gitaitis & Walcott, 2007).

Bacterial pathogens historically have been studied using culture-dependent methods (Jackson et al., 2013; Pereira et al., 2011). Sequence analysis of the entire 16S rRNA gene commonly is then used to identify the isolated (cultured) bacterial isolates to the genus and/or species level. However, these culturable bacteria comprise minor portions of most natural microbial communities (Ben-Dov et al., 2009; Rastogi & Sani, 2009). Microbial profiling of the V4 hypervariable region of the 16S rRNA gene (the primer pair used for the amplification of V4 region that is universal to bacteria with no bias) (Caporaso et al., 2011; Kozich et al., 2013), has been demonstrated to give reliable information about the composition of bacterial communities. The 16S rRNA profiling, a culture-independent method, can group bacteria based on similarities of the partial 16S rRNA sequences, with the groups referred to as operational taxonomic units (OTUs) (Tringe & Hugenholt, 2008; Větrovský & Baldrian, 2013). Operational taxonomic units are bacterial sequences that share a distinct level of similarity and

are used to measure the diversity of a bacterial community at a certain taxonomic level (Shade et al., 2013; Větrovský & Baldrian, 2013). Since many bacteria are non-culturable (Oliver, 2010), the use of culture-independent methods provides a more realistic view of the diversity of bacterial communities than culture-dependent methods (Jackson et al., 2013; Liu et al., 2012).

Culture-independent surveys conducted by Barret et al. (2015) revealed that seed-associated microbes from 28 seed samples representing a range of Brassicaceae were composed of 50-1000 bacterial and fungal OTUs. To our knowledge, the diversity of bacteria associated with onion seeds has not been examined using culture-independent methods, but could provide valuable insights into the complexity of bacterial communities associated with onion seeds, including plant pathogens, endophytes, human pathogens, and potentially beneficial bacteria. This study was performed to gain insight into the bacterial diversity associated with onion seed lots produced in South Africa by culturing the samples on agar media and by sequencing the V4 region of the 16S rRNA gene, using the Illumina MiSeq platform (Kozich et al., 2013; Schloss et al., 2009).

3.3 Materials and methods

Seed samples

Eighteen onion seed samples containing 20 gram of seeds per sample (two subsample were provided) were obtained from seed producers in the Northern and Western Cape Provinces, South Africa immediatedly after harvest (Table 1). The seeds were collected randomly from ten different seed production fields. Seven seed lots were from the Northern Cape and 11 seed lots were from the Western Cape. All seed lots were from a single cultivar that is grown widely across South Africa and is susceptible to bacterial blight. Of the 18 seed lots, three samples from the Western Cape had been treated with fungicides, namely Celest [fludioxonil (1ml/kg), Syngenta SA (Pty) Ltd], Thiram [thiocarbamate 5ml/kg) ARYSTA LifeScience South Africa] and a combination of Celest and Thiram [fludioxonil (1ml/kg) and thiocarbamate (5ml/kg), respectively]. These fungicide seed treatments are recommended for control of seed decay, damping-off and seedling blights caused by many seed- and soilborne fungi.

Culture-dependent method

Plating on agar media: Isolations from seed were carried out as described previously by Goszczynska et al. (2006), but samples were not surface-sterilised prior to isolations or DNA extraction. For each seed sample, two 5 g sub-samples were each crushed with a sterile pestle and mortar, and placed into separate sterile Erlenmeyer flasks containing 100 ml of a quarter-strength Ringer's buffer (Oxoid, Basingstoke, Hampshire, England). Flasks were then incubated for 30 min at 25°C on a rotary shaker at 180 rpm/min. For each sample, four 10-fold serial dilutions were made in a half-strength nutrient broth. An aliquot of 0.1 ml of each dilution was plated onto each of two petri dishes of tryptone glucose extract agar (TGA) medium (Difco). Plates were incubated at 25°C for five to seven days. After this time, total bacterial colony counts and the count of each colony type were performed and expressed as CFUs per gram (CFU/g) of seeds. All colony types isolated were purified on TGA plates using standard bacterial isolation and purification techniques, and incubated at 28°C for 72 h.

DNA extraction: Genomic DNA was extracted from each bacterial isolate using the cetyltrimethylammonium bromide (CTAB) method according to the DNA Miniprep protocol described by Wilson (1989). Colonies from each strain were picked from the 48 hours NA plates and transferred into microcentrifuge tubes with sterile STE (10mM Tris-HCl, 1mM EDTA, 2M NaCl, pH 8.0) buffer. Cell lysis was obtained by incubation with sodium dodecyl sulfate and selective precipitation of cell debris and polysaccharides with CTAB/NaCl. DNA was extracted with chloroform-isoamyl alcohol, precipitated with isopropanol, washed with ethanol, air dried and dissolved in 100 µl of nuclease free water. Purified DNA was quantified using a Dyna Quant 200 fluorometer (Hoefer, San Francisco, CA) and Hoescht H 33258 intercalating dye (Polysciences, Warrington, PA). Purified DNA was quantified using a NanoDrop (Inqaba Biotech). The DNA was stored at -20ºC until further analysis. PCR amplification of the 16S rRNA gene was performed with the universal primers fD1 (5'-AGA GTT TGA TCC TGG CTC AG-3') and rD1 (5'-AAG GAG GTG ATC CAG CCG CA-3') (Weisberg, 1981). The PCR reaction was prepared in a 25 µl volume containing 16.35 µ1 nuclease free water (Promega), 2.5 µl buffer, 2 µl MgCl2, 2 µl dNTPs; 0.1 µl forward and reverse primers, 0.15 µl Taq, and 1 µl DNA. PCR conditions were as follow: denaturation at 94ºC for 10 minutes; 30 cycles of denaturation at 94ºC for 1 min, annealing of primers at 55ºC for 1 minute, and elongation at 72ºC for 1 min; and a final extension for 5 minutes at 72ºC. The PCR products of ten different bacterial colonies (based on colony morphology) were selected randomly from the 18 seed lots for sequencing (Inqaba Biotechnology, South Africa) using

above mentioned primers as described by Weisburg (1981). The National center for Biotechnology Institute (NCBI) BLASTn (blast.ncbi.nlm.nih.gov) search was performed on the sequence and identity of the selected colonies were presented in Table 1 and the sequences were submitted to the Genbank (Table 1).

Culture-independent method

DNA extraction: Genomic DNA extraction was performed directly from each of the 18 onion seed lots. For each lot, 5 g of seed was ground in a mortar with a pestle, placed in a sterile plastic bottle and soaked with 20 ml of half strength Ringer's solution overnight. DNA isolation from cell immobilised on filter membrane was extracted using the phenol-chloroform protocol modified by Urakawa et al. (2010). Briefly, overnight cell suspension of each sample was filtered using the Sterivex filter membrane. The Sterivex filter membrater was removed using ethanol sterilised tweezer and cut aseptically using ethanol sterilised scalpel into several pieces and placed into a 2 ml bead-beating (lysing matrix E tube; MP Biomedicals), followed by addition of 300 µl of 2X TENS buffer (100 mM Tris-HCl [pH 8.0], 40 mM EDTA, 200 mM NaCl, 2% SDS) then vortexed vigorously. 1ml of phenol:chloroform:isoamyl alcohol (25:24:1) was added and then vortexed, followed by beads beating using Fastprep at setting 6 for 40 seconds. After bead beating, 2ml Heave Phase Locked Gel tube (Eppendorf, Westbury, NY) was pre-spinned at 14,000xg for 1 minutes, then centifuge for 10 min at 14,000xg to pellet bead. Aqueous phase was transferred to a 2.0-ml Phase Lock Gel tube, then 200 µl of 2X TENS was added. Bead beat on Fastprep at setting 6 for 40 seconds, centrifuged for 10 min at 12,500xg to pellet beads after that, 200 µl of 2X TENS was added. The beads were pelleted centrifuging for 10 min at 12,500xg; then 700 µl of phenol:chloroform:isoamyl alcohol (25:24:1) was added to the tube and mixed by reapeated inversion; and then then tube was centrifuged at 14,000xg for 5 min. Aqueous solution was transferred to 2 ml microcentrifuge tube, and then 600 µl of isopropanol was added to tube to precipipate nucleic acid from the aqueous; followed by addition of 6.0 µl of 15mg/ml Glycoblue and then mixed gently. The tube was incubated in a freezer at -80° for 10 min, and the centrigufed for 30 min at 12,000xg at room temperature. The supernatant was discarded (make sure that the blue pellet remain in the tube). The tube containing the blue pellet was precipitated with 1 ml of 80% ethanol, and then centriguged for 30 min at 12,000xg. Discarded alcohol and air dried for 5-10 min. Re-suspended sample in 50 µl nuclease free water and stored at -20°C until further analysis.

Amplicon library construction and sequencing: Amplicon libraries were constructed following two rounds of PCR amplification. The first amplification was performed with the PCR primers 515 and 806 (Caporaso et al., 2011), which target the V4 region of the 16S rRNA gene. The PCR reaction was performed with a high-fidelity polymerase (AccuPrime Taq DNA Polymerase System; Invitrogen) following the manufacturer's protocol. Cycling conditions were: 95°C for 2 min; followed by 30 cycles of amplification at 95°C for 20 s, 55°C for 15 s and 72°C for 5 min; with a final extension step at 72°C for 10 min. All amplicons were purified with the Agencourt AMpure XP system (Beckman Coulter, USA) and quantified with QuantIT PicoGreen (Invitrogen). The second round of amplification was performed with 5 µl of purified amplicons and the primers containing the Illumina adapter and indexes (Caporaso et al., 2011). PCR cycling conditions were: 94°C for 2 min; followed by 12 cycles of amplification at 94°C for 1 min, 55°C for 1 min, and 68°C for 1 min; and a final extension step at 68°C for 10 min. All amplicons were purified and quantified as previously described. The purified amplicons were pooled in equimolar concentrations and the final concentration of the library was determined using a qPCR Next Generation Sequencing (NGS) Library Quantification Kit (Kapa Biosystems). Amplicon libraries were mixed with 10% Phix control according to the Illumina protocol. Sequencing runs were performed with MiSeq Reagent Kit v2 (500 cycles) as described by Caporaso et al. (2011).

Microbial profiling data analysis

Assessment of microbial diversity: Standard operating procedure (SOP) is a procedure used by the Schloss lab to process their 16S rRNA gene input data generated by Illumina's MiSeq platform (Barret et al., 2015; Kozich et al., 2013; http://www.mothur.org/wiki/MiSeq_SOP). Raw reads were obtained from MiSeq as pairs of fastq files with each pair representing the two sets of reads per sample (for example V1_S194_L001_R1-001.fastq and V1_S194_L001_R2-001.fastq) (Barret et al., 2015; Kozich et al., 2013; http://www.mothur.org/wiki/MiSeq_SOP). Fastq files contain both the input data and the quality score data. Because the reads are about 250 bp in length, the sequencing was done from either end of each fragment, this results in a significant overlap between the forward and reverse reads in each pair (Kozich et al., 2013; http://www.mothur.org/wiki/MiSeq_SOP). The pairs of reads were combined into contigs, the make.contigs tool creates the contigs and uses the paired collection as input (Kozich et al., 2013; http://www.mothur.org/wiki/MiSeq_SOP).

Unique.seq tool was used to determne the unique reads and the number these different reads observed were recorded in the original dataset because community samples normally contain a large numbers of the same organism (http://www.mothur.org/wiki/MiSeq_SOP).

The sequences were aligned using the 16S rRNA gene SILVA alignment tool (Quast et al., 2013). All the sequences that did not align correctly were removed from the data set. Chimeric sequences were detected using UCHIME (Edgar et al., 2011) and then removed from the data set. Taxonomic affiliations were performed with a Bayesian classifier (Wang et al., 2007) using the Greengenes reference database (DeSantis et al., 2006; McDonald et al., 2012). Non-classified sequences (0.001% of the 16S rRNA gene sequences) or sequences belonging to the Archaea or Eukaryota chloroplast or mitochondria were discarded (Barret et al., 2015; Kozich et al., 2013; http://www.mothur.org/wiki/MiSeq_SOP). Finally, sequences were divided into groups corresponding to their taxonomic rank (level of order) and then assigned to OTUs Operational taxonomic units (OTUs). Operational taxonomic units were defined as aOTUs at a threshold of ≥ 0.03% of the library size at 97% identity cut-off (Kozich et al., 2013; http://www.mothur.org/wiki/MiSeq_SOP).

All the statistical analysis for bacterial community chapter were performed using Microsoft Excel 2016. After removal of unwanted taxon step performed using Mothur, a file containing taxonomy level (phylum, family, and genus) was dragged in Excel (Microsoft Excel 2016). The spreadsheet created by Mothur was located and opened in Excel to make a bar chart. All the data that must be included in the bar chart were selected, the column and row headers were included in order for them to become the labels in the bar chart. Appropriate header cells were typed for different labels. The chart wizard toolbar button was then clicked or chart from the insert menu was chosen. Column was selected under chart type (it is the default setting)., a subtype of a bar graph was selected from the display on the right. The titles of the chart for the X and Y-axis were entered; these titles will appear in the appropriate places on the bar graph. A bar graph was copied and inserted in a PowerPoint slide and then the chart toolbar was used to make any final adjustments on the bar graph.

Diversity of bacterial community in this study was analysed using a heatmap which was constructed using XLStat (https://www.xlstat.com/). A heatmap is a graphical representation of data that uses a system of color-coding to represent different values. XLStat was used to construct a heatmap as follow: In the general tab, the data matrix in the features/individuals table field was selected. The individuals were represented by the samples. Nucleotide

sequences were stored in rows in the dataset. The non-specific filtering option were activated in the options tab and interquartile range < (In statistic interquartile range is defined as midspread or middle 50%) was selected and a threshold of 0.25 was entered. All nucleotide sequences will be eliminated with an interquartile range lower than 0.3 (for example, with low inconsistency).

The chart size was optimised by playing with the width and height and the colour scale of the heatmap was selected in the chart tab. Abundant OTUs (aOTUs) representing at least 0.1% of the library size were used for microbial community analysis (Barret et al., 2015). In figure 3.2, the seed lots are listed on the right vertical axis, and OTUs on the x-axis, red colour represent the most relations of the OTUs and the least OTUs were represented by blue colour. The dendrogram along the sides of the heatmap shows how the variables and the rows are clustered independently.

3.4 Results

Culture-dependent method

The number of cultivable bacteria recovered from TGA plates for both treated and non-treated seed samples ranged from 5 x 10^3 to 6.2 x 10^6 CFU/g seed. Cultivable isolates belonged to the phylum Proteobacteria and included representatives of *Acinetobacteria*. *Enterobacter*, *Erwinia*, *Microbacterium*, *Pantoea* and *Pseudomonas* (Table 1). The bacteria of the genus *Pantoea* were isolated from all 18 seed lots at populations ranging from 6.0 x 10^3 to 6.7 x 10^6 CFU/g seed, depending on the sample. *Acinetobacteria* were isolated from three seed lots, V7 and V10, at a range of 5.0 x 10^3 to 4.0 x 10^4 CFU/g seed; and *Pseudomonas* was isolated from two seed lots, V5 at 7.0 x 10^4 CFU/g seed and V10 at 8.0 x 10^4 CFU/g seed. *Enterobacter*, *Erwinia* and *Microbacterium* were isolated from only one seed lot each, V10, V16 and V7, respectively, at populations of 6.0 x 10^4, 6.0 x 10^3 and 3.0 x 10^4 CFU/g seed, respectively (Table 1). Sequences of the bacterial isolates were deposited in GenBank, accession numbers MF138095 to MF138104 (Table 1).

Culture-independent method

A total of 587141 sequence reads was generated from the 18 onion seed lots, with the number of filtered sequence reads at 159187 (Table 3.3). For all 18 seed lots, a total of 512 OTUs (Table S1) were detected, representing six phyla (Fig. 3.1a), 14 families (Fig. 3.1b) and eight

genera that could be identified, namely, *Acinetobacter*, *Enterococcus*, *Pantoea*, *Pseudomonas*, *Providentia*, *Sphingobacterium*, *Sphingomonas* and *Stenotrophomonas* (Fig. 3.1c). The number of OTUs detected in each seed sample was used as a simple measure of bacterial community diversity. An average of 62 OTUs was detected across all 18 seed samples (Fig. 3.2, Table S1), and ranged from 1 to 3202 for individual seed samples. Of the 512 OTUs recognised, 233 (45.5%) were represented by just one sequence in a single seed sample, and 93 (18.2%) were represented by only two sequences. Unexpectedly, the seed lot treated with fungicide Thiram had the most OTUs (3397), followed by the non-treated seed lots V10 (3268), V5 (1863) and a seed lot treated with the fungicide Celest (1861). The non-treated seed lots V1, V3, V6, V8 and V14 had few OTUs, ranging from 15 to 34.

The phylum Proteobacteria dominated the bacterial seed community in all 18 seed lots, with a relative abundance of 96% (Fig. 3.1a). Bacteroidetes, Firmicutes, Planctomycetes, Actinobateria, and Acidobacteria represented 1% or less of the sequences that could be identified to phylum. The Enterobacteriaceae dominated the family-level of OTUs detected in the 18 seed lots (Fig. 3.1b), of which the bacterial sequences classified as *Pantoea* were present in the seed lots (Fig. 3.1c). The genera *Acinetobacter* and *Pseudomonas* were detected in 13 and 15 of the seed lots, respectively (Table 3.2). *Sphingobacterium* was detected in eight seed lots, while *Sphingomonas* was detected in 11 seed lots. *Enterococcus*, a genus of lactic acid bacteria in the phylum Firmicutes, was detected in five seed lots. *Providential* and *Stenotrophomonas* were each detected in only four seed lots.

Diversity indices of the OTUs (Fig. 3.2) illustrated the dominant bacteria (e.g., phyla, family, and genus) associated with the 18 onion seed lots. The heatmap (Fig. 3.2) showed that OTU3, OTU4 and OTU7 (Table S1) were the most abundant among the 18 seed lots. OTU3 was detected in all seed lots except V1 and V6. OTU4 was not detected in five of the seed lots, V1, V8, V9, V13 and V14. OTU147 was detected in seed lot V1 and V6, while OTU271 was identified only in seed lot V7.

3.5 Discussion

Microbial communities in and on seeds, may have useful or harmful effects on plant growth and fitness. In this study, culture-dependent and culture-independent methods were used to analyse the bacterial diversity associated with 18 seed lots of a single cultivar produced in two provinces in South Africa. *Pantoea* was the dominant genus detected among the 18 seed lots with both methods. Other bacteria detected using both methods are known to be seed endophytes, e.g., *Pseudomonas* and *Enterobacter* (Johnston-Monje & Raizada, 2011; Rijavec et al., 2007). Enterobacteriaceae and Pseudomonadeae have been isolated by other research groups from seed of wheat, canola, spinach and broccoli (Links et al., 2014; Lopez-Velasco, 2013; Weiss et al., 2007).

Seeds can harbour a diversity of bacterial endophytes (Coombs & Franco, 2003; Johnston-Monje & Raizada, 2011; Truyens et al., 2015). Seed endophytes are readily seed transmitted, i.e., transmitted from generation to generation of plants that are propagated by seed (Truyens et al., 2015). Bacterial seed endophytes have been detected using both culture-dependent and culture-independent methods in plant tissues (Hardoim et al., 2012; Johnston & Raizada et al., 2011; Links et al., 2014; Liu et al., 2012; Lopez-Velasco et al., 2013; Truyens et al., 2015). In this study, the samples of onion seeds from which bacterial communities were characterised were not surface-sterilized prior to isolations or DNA extraction, which prevented distinction of bacteria present on the seed vs. those present within the seed. Attachment of bacteria to the seed surface (coat) allows bacterial pathogens to evade the plant defence mechanisms within the seed, with the results that the seeds become a passive carrier of the pathogens.

The use of culture independent methods to examine microbial communities in seeds can provide information on diversity of the bacterial communities on and within seed (Hardoim et al., 2012; Liu et al., 2012; Lopez-Velasco et al., 2013). In this study, the culture independent method based on the V4 region of 16S rRNA gene sequences, detected few seed endophytes such as *Stenotrophomonas* and *Sphingobacterium*, that were not isolated using the culture-based method. For example, *Stenotrophomonas* was present only in seed lots V9, V16, seed lot coated with Thiram and Celeste, respectively. Hardoim et al. (2012) used culture dependent and culture independent methods in their study to evaluate the endophytic bacterial community of surface sterilised rice seeds, counting two sequential rice groups. The authors identified nine bacterial genera which include *Stenotrophomonas* from both seed groups.

Oliver (2010) reported that culture-independent methods of microbial analyses may suffer from primer bias or may be affected by various lysis of different bacterial species, resulting in skewed representation of the bacterial species identified. These methods do, however, provide comprehensive data of the bacterial community present, including the DNA of viable but non-culturable bacteria (Pereira et al., 2011). For example, the genera *Providencia* and *Sphingobacterium* did not grow on medium used in this study. These genera were detected only by the culture-independent method. One limitation with this method is that it often fails to resolve bacterial identification to the species level and is commonly inadequate to taxonomic association at the genus level (Barret et al., 2015; Větrovský & Baldrian, 2013).

Seed treatments have many applications in agriculture. However, the three seed lots treated with fungicides in this study did not differ significantly from the 15 non-treated seed lots in terms of bacterial community diversity. These results suggest that neither Celest nor Thiram affected the structure of the seed-associated bacterial communities on these seed lots. Thiram and Celest are the most commonly used protectant fungicides on onion seed in South Africa. They reduce losses from seedling blights caused by seed borne and soil borne fungal pathogens, but not bacterial plant pathogens (www.nunhemsusa.com; Syngenta, South Africa). Therefore, the three onion seed lots treated with these fungicides were not expected to differ in bacterial community diversity.

Overall, the culture-dependent and -independent analyses used in this study indicated that the majority of bacteria associated with fungicide-treated and non-treated onion seed lots were members of the phylum Proteobacteria. Although we used seed lots of a single cultivar in the study, common bacterial genera were detected in association with all 18 onion seed lots evaluated, which were similar to the plant and common soilborne bacterial pathogens and endophytes reported previously on seeds of rice, wheat, and maize (Johnston & Raizade, 2011; Liu et al., 2012; Truyens et al., 2015). Among the detected bacterial genera (*Pantoea*, *Pseudomonas*, *Acinetobacteria*, *Sphingomonas* and *Enterobacter*), some have been reported to be plant growth promoters and biological control agents (Link et al., 2014; Liu et al., 2012). For example, *P. agglomerans* 2066-7 was reported as a biological control agent of bacterial onion diseases such as *Pseudomonas marginalis*, *P. ananatis*, *P. viridiflava* and *Xanthomonas retroflexus* (Sadik et al., 2013). Links et al. (2014) reported that some strains of *P. agglomerans* are used as biological control agents that may be important in protecting seed from fungal pathogens such as *Altanaria* sp. Puente et al. (2009) also demonstrated that the seedborne

endophytes *Pseudomonas* sp. SENDO 2 and *Acinetobacter* sp. SENDO 1 improved seedling growth and growth of cardon cactus. In addition, researchers demonstrated that, *Enterobacter cloacae* strain 501R3 plays a significant role in the preventing of damping-off caused by *Pythium ultimum* through competitive colonisation of the rhizosphere soils (Hardoim et al., 2012; Kageyama & Nelson, 2003; Roberts et al., 2007).

In general, the majority of studies rely on analysis of 16S rRNA gene sequences to survey bacterial community diversity. Given the state of the results presented in this study, the V4 region of the 16S rRNA gene sequences generated by a high-throughput deep-sequencing technologies will usually classify organisms to the genus level. The culture-independent approach allowed for the identification of several OTUs present in limited abundance that were not identified using the culture-dependent method. However, a single 16S rRNA OTU can have large amount of species which remains largely unexplored in studies that are confirmed to even the most highly variable regions of 16S rRNA gene. Since this study is the first to determine the bacterial community in and on onion seeds, we proposed that different cultivars should be evaluated to determine if different onion cultivars contain similar bacterial communities. Different genes are required to resolve the diversity within certain taxonomic groups because bacterial phyla differ in their genomic contents. In this case, a species-specific bacterial marker based on a fragment of gyrase B (*gyrB*) gene should be used for determining microbial diversity in and on onion seeds. The *gyrB* gene is routinely used for bacterial identification, phylogenetic and in the study of bacterial diversity. Furthermore, it contains at least two highly conserved regions that are suitable for designing low degeneracy primers and are separated by distance appropriate to high-throughput sequencing platform (in the case of Illumina MiSeq).

3.6 References

Adam, E., Bernhart, M., Müller, H., Winkler, J. & Berg, G. 2016. The *Cucurbita pepo* seed microbiome: genotype-specific composition and implications for breeding. *Plant and Soil* 422:35-39.

Baker, K.F. & Smith, S.H. 1966. Dynamics of seed transmission of plant pathogens. *Annu Rev Phytopathol* 4:311-332.

Barret, M., Briand, M., Bonneau, S., Préveaux, A., Valière, S., Bouchez, O., Hunault, G., Simoneau, P. & Jacques, M.A. 2015. Emergence shapes the structure of the seed microbiota. *Appl Environ Microbiol* 81:1257-1266.

Barret, M., Guimbaud, J.F., Darrasse, A. & Jacques, M.A. 2016. Plant microbiota affects seed transmission of phytopathogenic microorganisms. *Mol Plant Pathol* 17:791-795.

Ben-Dov, E., Kramarsky-Winter, E. & Kushmaro, A. 2009. An in situ method for cultivating microorganisms using a double encapsulation technique. FEMS *Microbiol Ecol* 68:363-371.

Benkeblia, N. & Lanzotti, V. 2007. *Allium* Thiosulfinates: Chemistry, biological properties and their potential utilization in food preservation. *Food* 1:193-201.

Bulgarelli, D., Schlaeppi, K., Spaepen, S., Ver Loren van Themaat, E. & Schulze-Lefert, P. 2013. Structure and functions of the bacterial microbiota of plants. *Annu Rev Plant Biol* 64:807-838.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh P.J., Fierer, N. & Knight, R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci USA* 108:4516-22.

Coombs, J.T. & Franco, C.M.M. 2003. Visualization of an endophytic *Streptomyces* species in wheat seed. *Appl Environ Microbiol* 69:4260-4262.

Department of Agriculture, Forestry and Fisheries, 2015. A profile of the South African onion market value chain. Department of Agriculture, Pretoria, South Africa.

Darrasse, A., Darsonval, A., Boureau, T., Brisset, M.N., Durand, K. & Jacques, M.A. 2010 Transmission of plant-pathogenic bacteria by nonhost seeds without induction of an associated defence reaction at emergence. *Appl Environ Microbiol* 76:6787-6796.

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P. & Andersen, G.L. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *AEM* 72:5069-5072.

Dutta, B., Gitaitis, R., Smith, S. & Langston Jr., D. 2014. Interactions of seedborne bacterial

pathogens with host and non-host plants in relation to seed infestation and seedlingtransmission. *PLoS ONE* 9(6): e99215.

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

Gitaitis, R.D. & Walcott, R.R. 2007. The epidemiology and management of seedborne bacterial diseases. *Annu Rev Phytopathol* 45:371-397.

Goszczynska, T., Moloto, V.M., Venter, S.N. & Coutinho, T.A. 2006. Isolation and identification of *Pantoea ananatis* from onion seed in South Africa. *Seed Science Technol* 34:677-690.

Guan, K.L. 2009. Seed physiological ecology. Chinese Agricultural Press, Beijing, pp 1-7 (in Chinese).

Hardoim, P.R., Hardoim, C.C.P., van Overbeek, L.S. & van Elsas, J.D. 2012. Dynamics of seedborne rice endophytes on early plant growth stages. *PLoS ONE* 7(2): e30438.

Ibekwe, A. & Grieve, C. 2004. Changes in developing plant microbial community structure as affected by contaminated water. *FEMS Microbiol. Lett* 48:239-248.

Jackson, C.R., Randolph, K.C., Osborn S.L. & Heather, L. & Tyler, H.L. 2013. Culture dependent and independent analysis of bacterial communities associated with commercial salad leaf vegetables. *BMC Microbiol* 13:274-285.

Johnston-Monje, D., Lundberg, D.S., Lazarovits, G., Reis, V.M. & Raizada, M.N. 2016. Bacterial populations in juvenile maize rhizospheres originate from both seed and soil. *Plant Soil* 405:337-355.

Johnston-Monje, D. & Raizada, M.N. 2011. Conservation and diversity of seed associated endophytes in *Zea* across boundaries of evolution, ethnography and ecology. *PLoS ONE* 6(6): e20396.

Kageyama, K. & Nelson, E.B. 2003. Differential inactivation of seed exudate stimulation of *Pythium ultimum* sporangium germination by Enterobacter cloacae influences control efficacy on different plant species. *Appl Environ Microbiol* 69:1114-1120.

Klaedtke, S., Jacques, M-A., Raggi, L., Préveaux, A., Bonneau, S., Negri, V., Chable, V. & Barret, M. 2015. Terroir is a key driver of seed-associated microbial assemblages. *Environ Microbiol* 18:1792-1804.

Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K. & Schloss, P.D. 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* 79:5112-5120.

Links, M.G., Demeke, T., Gräfenhan, T., Hill, J.E., Hemmingsen, S.M. & Dumonceaux, T.J. 2014. Simultaneous profiling of seed-associated bacteria and fungi reveals antagonistic interactions between microorganisms within a shared epiphytic microbiome on *Triticum* and *Brassica* seeds. *New Phytol* 202:542-553.

Liu, Y., Zuo, S., Liwen Xu, L., Zou, Y. & Song, W. 2012. Study on diversity of endophytic bacterial communities in seeds of hybrid maize and their parental lines. *Arch Microbiol* 194:1001-1012.

Lopez-Velasco, G., Carder, P.A., Welbaum, G.E. & Ponder, M.A. 2013. Diversity of the spinach (*Spinacia oleracea*) spermosphere and phyllosphere bacterial communities. *FEMS Microbiol Lett* 346:146-154.

Lozupone, C.A. & Knight, R. 2005. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* 71:8228-8235.

Maude, R.B. 1996. Seedborne diseases and their control: principles and practice. CAB International, Oxon, United Kingdom.

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

Nelson, E.B. 2004. Microbial dynamics and interactions in the spermosphere. *Ann Rev Phytopathol* 42:271-309.

Oliver, J.D. 2010. Recent findings on the viable but nonculturable state in pathogenic bacteria. *FEMS Microbiol Rev* 34:415-425.

van Overbeek, L.S, Franke, A.C., Nijhuis, E.H., Groeneveld, R.M., da Rocha, U.N. & Lotz, L.A. 2011. Bacterial communities associated with *Chenopodium album* and *Stellaria* media seeds from arable soils. *Microb Ecol* 62:257-264.

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. *Int Sch Res Netwk Ecol* 2011:1-10.

Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W., Peplies, J. & Glockner, F.O. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* 35:7188-7196.

Puente, M.E., Li, C.Y. & Bashan Y. 2009. Endophytic bacteria in cacti seeds can improve the development of cactus seedlings. *Environ Exp Bot* 66:402-408.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. & Glöckner F.O. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41:590-596.

Rastogi G. & Sani, R.K. 2011. Molecular Techniques to Assess Microbial Community Structure, Function, and Dynamics in the *Environ Micro Microb Technol* pp 29-57.

Rezki, S., Campion, C., Iacomi-Vasilescu, B.I., Preveaux1, A., Toualbia, Y., Bonneau, S., Briand, M., Laurent, E., Hunault, G., Simoneau, P., Jacques, M-A. & Barret, M. 2016. Differences in stability of seed-associated microbial assemblages in response to invasion by phytopathogenic microorganisms. *PeerJ* 4:e1923.

Rijavec, T., Lapanje, A., Dermastia, M. & Rupnik, M. 2007. Isolation of bacterial endophytes from germinated maize kernels. *Can J Microbiol* 53:802-808.

Roberts, D.P., McKenna, L.F., Lohrke, S.M., Rehner, S. & de Souza. J.T. 2007. Pyruvate dehydrogenase activity is important for colonization of seeds and roots by Enterobacter cloacae. *Soil Biol Biochem* 39:2150-2159.

Sadik, S., Mazouz, H., Bouaichi, A. Benbouazza, A. & Achbani, H.E. 2013. Biological Control of Bacterial Onion Diseases using a Bacterium, *Pantoea agglomerans* 2066-7. *IJSR* 2319-7064.

SANSOR Annual Report 2013/2014. 2015. South African National Seed Organisation, South Africa, Pretoria.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J. & Weber, C.F. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75:7537-7541.

Shade, A., McManus, P.S. & Handelsman, J. 2013. Unexpected diversity during community succession in the apple flower microbiome. *mBio* 4:e00602e00612.

Spinelli, F., Ciampolini, F., Cresti, M., Geider, K. & Costa, G. 2005. Influence of stigmatic morphology on flower colonization by *Erwinia amylovora* and *Pantoea agglomerans*. *Eur J Plant Pathol* 13:395-405.

Terrasson, E., Darrasse, A., Righetti, K., Buitink, J., Lalanne, D., Ly Vu, B., Pelletier, S., Bolingue, W., Jacques, M.A. & Leprince, O. 2015. Identification of a molecular dialogue between developing seeds of Medicago truncatula and seedborne xanthomonads. *J Exp Bot* 66:3737-52.

The National Agricultural Directory, 2011. Field crops and horticulture. Vegetables.

Department of Agriculture, Pretoria, South Africa.

Tringe, S.G. & Hugenholtz, P. 2008. A renaissance for the pioneering 16S rRNA gene. *Curr Opin Microbiol* 11:442-446.

Truyens, S., Weyens, N., Cuypers, A. & Vangronsveld, J. 2015. Bacterial seed endophytes: genera, vertical transmission and interaction with plants. *Environ Microbiol Rep* 7:40-50.

Urakawa, H., Martens-Habbena, W. & Stahl, D.A. 2010. High abundance of ammonia-oxidizing Archaea in coastal waters, determined using a modified DNA extraction method. *Appl Environ Microbiol* 76:2129-2135.

Van Der Meer, Q.P. 1997. Old and new crops within edible *allium*. Acta Hort. 433:17-31.

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, Q., Garrity, G.M., Tiedje, J.M. & Cole, J.R. 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 73:5261-5267.

Weisburg, W.G., Barns, S.M., Pelletier, D.A. & Lane, D.J. 1991. 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* 173:697-703.

Weiss, A., Hertel, C., Grothe, S., Ha, D. & Hammes, W.P. 2007. Characterization of the cultivable microbiota of sprouts and their potential for application as protective cultures. *Syst Appl Microbiol* 30:483-493.

Wilson, K. 1989. Preparation of genomic DNA from bacteria. In Current protocols in molecular biology, (eds. F.M. Ausubel, R. Brent, R.E. Kimgston, D.D. Moore, J.G. Seidman, J.A. Smith and K. Struhl), 241-245, John Wiley & Sons, Inc, New York, N.Y.

3.7 Tables

Table 3.1 Bacterial genera detected in association with onion seed lots produced in South Africa, based on isolations using a culture-dependent method of plating a crushed extract of a 5 g sample of each seed lot onto tryptone glucose extract agar.

Seed lot	Province in South Africa where the seed was produced	Production field	Genus and species[a]	CFU/g of seed	GenBank accession number
V1	Western Province	F1	*P. agglomerans*	2.9×10^5	
V2	Western Province	F2	*P. agglomerans*	8.0×10^6	
V3	Western Province	F2	*P. agglomerans*	1.7×10^5	
V4	Western Province	F3	*P. agglomerans*	1.6×10^5	
V5	Western Province	F3	*P. agglomerans*	8.2×10^5	
			P. syringae	7.0×10^4	MF138095
V6	Western Province	F4	*P. agglomerans*	3.6×10^4	
V7	Northern Province	F5	*P. agglomerans*	1.6×10^5	
			Acinetobacter sp.	5.0×10^3	MF138097 & MF138098
			Microbacterium sp.	3.0×10^4	MF138104
V8	Northern Province	F6	*P. agglomerans*	6.7×10^6	
V9	Northern Province	F6	*P. agglomerans*	4.0×10^4	
V10	Northern Province	F6	*P. agglomerans*	$6. 5 \times 10^5$	
			Acinetobacter sp.	4.0×10^4	MF138099 & MF138100
			Enterobacter sp.	6.0×10^4	MF138101
			P. syringae	8.0×10^4	MF138096

V12	Northern Province	F7	*P. agglomerans*	1.4×10^5	
V13	Northern Province	F8	*P. agglomerans*	6.1×10^4	
V14	Northern Province	F9	*P. agglomerans*	1.4×10^5	
V16	Western Province	F10	*P. agglomerans*	5.0×10^4	
			Erwinia sp.	6.0×10^3	MF138102
Thiram[b]	Western Province	F10	*P. agglomerans*	5.2×10^5	
Celest[b]	Western Province	F10	*P. agglomerans*	2.0×10^4	
Cel_Thiram[b]	Western Province	F10	*P. agglomerans*	3.1×10^5	
Old SL	Western Province	F10	*P. agglomerans*	4.3×10^5	
			Erwinia	2.0×10^4	MF138103

[a] *P. agglomeras* and *P. syringae* were identified to species level based on procedures used such as indole production to identify *P. agglomerans* and carbon source utilization to identify *P. syringae*. Other genera we not identified to species because of the lack of a quick identification procedure for species identification in our laboratory.

[b] Treated seed lots are listed with the name of the fungicide treatment. All other seed lots evaluated were not treated with fungicides. Thiram contains the active ingredient thiocarbamate (750g/kg). Celest contains the active ingredient fludioxonil (1ml/kgl), Cel_Thiram contains both the ingredients fludioxonil (1ml/kg) and thiocarbamate (5ml/kg). Old SL = Seed lot that was harvested three years before this study.

Table 3.2 Bacterial sequences identified from onion seed lots using the 16S rRNA gene, and relative abundance (percentages) of these sequences.

Seed lot	Province in South Africa	Production field[a]	Phylum	Genus	Relative abundance (%)[b]
V1	Western Cape	F1	Proteobacteria	*Acinetobacter*	0.79
				Sphingomonas	0.39
				Pantoea	13.83
				Providencia	2.77
			Bacteriodetes	*Sphingobacterium*	1.98
V2	Western Cape	F2	Proteobacteria	*Acinetobacter*	0.04
				Pantoea	27.97
				Pseudomonas	1.64
				Sphingomonas	0.01
V3	Western Cape	F2	Proteobacteria	*Pantoea*	27.02
				Pseudomonas	0.11
V4	Western Cape	F3	Proteobacteria	*Acinetobacter*	0.10
				Pantoea	5.72
				Pseudomonas	25.54
				Sphingomonas	0.10
V5	Western Cape	F3	Proteobacteria	*Pantoea*	4.41
				Pseudomonas	28.75

			Bacteriodetes	*Sphingobacterium*	0.02
V6	Western Cape	F4	Proteobacteria	*Pantoea*	1.34
V7	Northerm Cape	F5	Proteobacteria	*Acinetobacter*	0.01
				Pantoea	57.47
				Pseudomonas	2.82
				Sphingomonas	0.01
			Bacteriodetes	*Sphingobacterium*	0.01
			Firmicutes	*Enterococcus*	0.05
V8	Northerm Cape	F6	Proteobacteria	*Acinetobacter*	5.45
			Firmicutes	*Enterococcus*	1.21
				Pantoea	11.52
				Pseudomonas	7.88
			Bacteriodetes	*Sphingobacterium*	1.21
				Sphingomonas	1.82
V9	Northern Cape	F6	Proteobacteria	*Acinetobacter*	1.68
				Pantoea	6.80
				Providentia	0.20
				Pseudomonas	22.00
				Stenotrophomonas	0.39
				Sphingomonas	0.10
			Bacteriodetes	*Sphingobacterium*	0.89

V10	Northern Cape	F6	Proteobacteria	*Acinetobacter*	0.03
				Pantoea	22.34
				Pseudomonas	25.92
				Sphingomonas	0.02
			Firmicutes	*Enterococcus*	0.12
V12	Northern cape	F7	Proteobacteria	*Acinetobacter*	0.07
				Pantoea	21.29
				Pseudomonas	0.59
V13	Northern Cape	F8	Proteobacteria	*Acinetobacter*	3.21
				Pantoea	28.11
				Pseudomonas	11.24
V14	Northern Cape	F9	Proteobacteria	*Acinetobacter*	0.85
				Pantoea	62.00
				Pseudomonas	0.14
			Bacteriodetes	*Sphingobacterium*	0.14
V16	Western Cape	F10	Proteobacteria	*Acinetobacter*	0.04
				Pantoea	30.39
				Providentia	0.09
				Pseudomonas	4.18
				Stenotrophomonas	1.45
			Firmicutes	*Enterococcus*	

Thiram	Western Cape	F10	Proteobacteria	*Acinetobacteria*	0.07
				Pantoea	6.11
				Sphingomonas	0.04
				Stenotrophomonas	0.01
Celest	Western Cape	F10	Proteobacteria	*Acinetobacteria*	0.01
				Pantoea	19.81
				Pseudomonas	0.95
				Providentia	0.03
				Sphingomonas	0.004
				Stenotrophomonas	0.28
			Bacteriodetes	*Sphingobacterium*	0.19
			Firmicutes	*Enterococcus*	0.004
Thi_Cel	Western Cape	F10	Proteobacteria	*Pantoea*	9.93
				Pseudomonas	4.67
				Sphingomonas	0.01
Old_SL	Western Cape	F10	Proteobacteria	*Pantoea*	11.55
				Pseudomonas	0.14
				Sphingomonas	0.03

[a] F = seed production field.

[b] Relative abundance (%) was calculated as: (how common or rare the organism is observed in a seed lot/total of all organisms present in that seed lot)*100. For example, *Pantoea* was observed 35 times in seed lot V1 and the number of organisms detected in seed lot V1 is 253. Relative abundance (%) of *Pantoea* for V1= (35/253)*100 which is equal to 13.83.

Table 3.3 Diversity statistics for bacterial communities associated with samples of each of 18 onion seed lots of the same cultivar collected from seed producers in Northern and Western Cape Provinces of South Africa, with the communities based on sequencing of the 16S rRNA DNA sequences detected in crushed extracts of a sample of each seed lot.

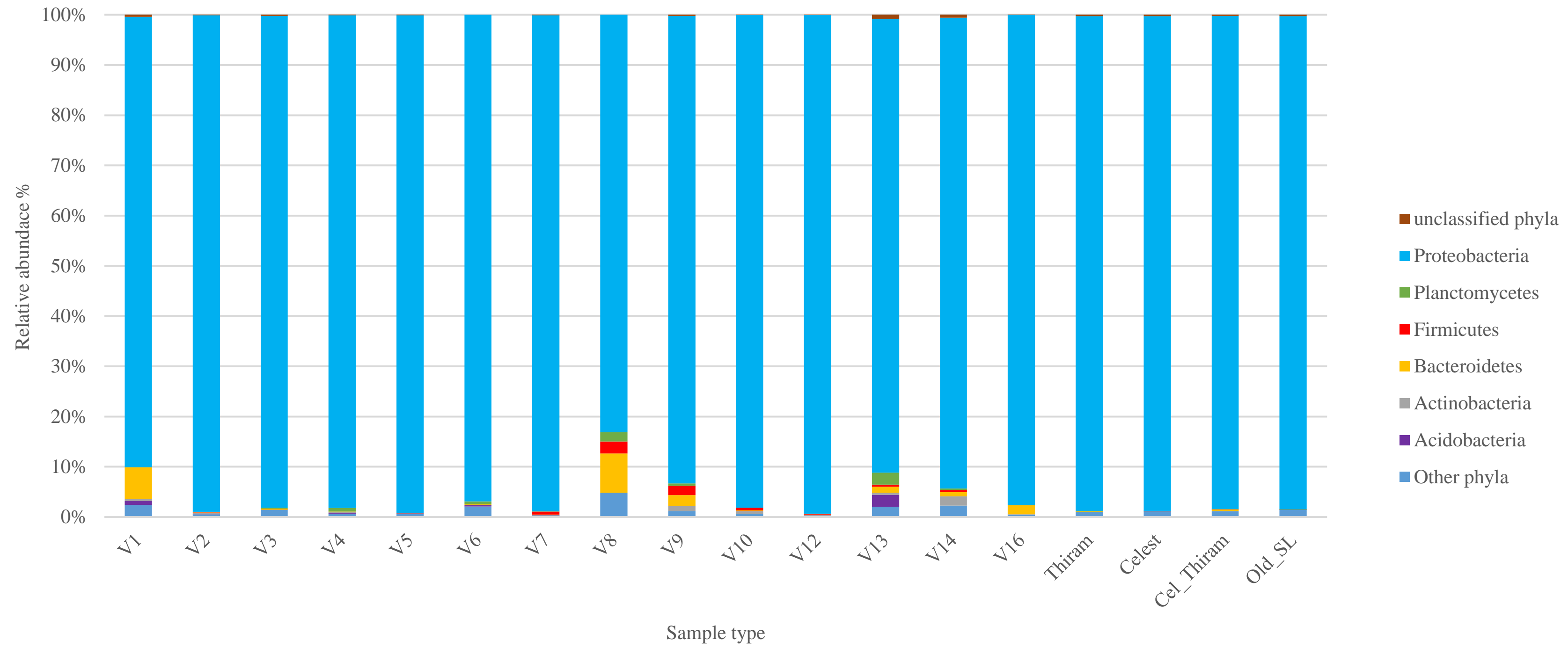
Seed lot[a]	Raw sequencing reads	Processed (filtered) sequencing reads	Number of OTUs (97% similarity)
V1	1524	253	29
V2	28111	7237	57
V3	7382	1910	39
V4	4119	1014	27
V5	20484	6282	43
V6	4326	674	29
V7	38333	15100	80
V8	1109	165	35
V9	4638	1014	80
V10	34801	12348	79
V11	32580	8216	61
V12	1140	249	26
V13	2144	705	41
V14	53722	15916	103
V15	92416	25234	113

V16	41429	11736	103
V17	117689	29667	92
Old_SL	101194	21467	80

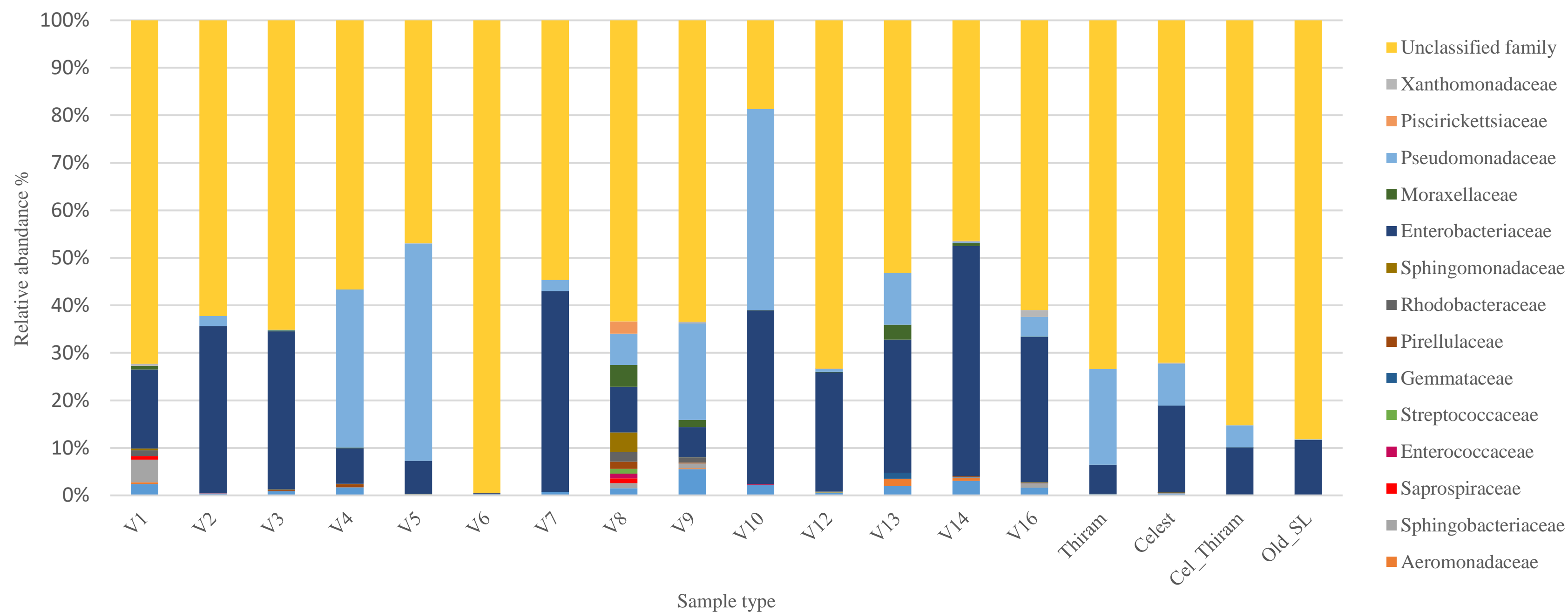
[a] Refer to Table 1 for details of the seed lots.

3.8 Figures

1a)



1b)



1c)

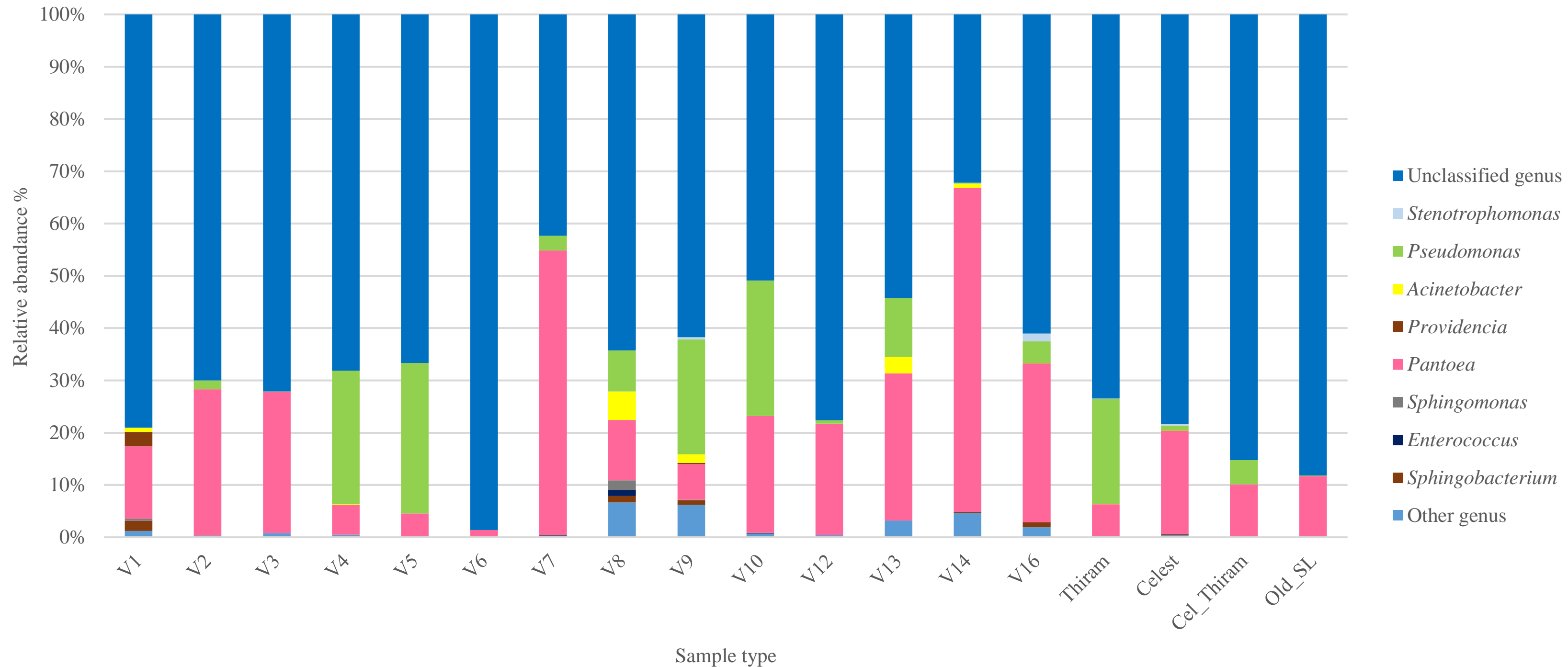


Figure 3.1 Relative abundance of bacterial sequences of the V4 region of the 16S rRNA gene detected from DNA extracted from 18 onion seed lots of the same cultivar grown in two provinces in South Africa. Results are represented as relative abundance of the total bacteria classified to phylum (a), family (b) and genus (c). The percentage of sequences reads was calculated based on the total number of reads classified beyond the bacterial order using the Greengene taxonomic database. Orders with <1% relative abundance in any sample type were identified as other (phyla, families and genera). "Unclassified" refers to all bacterial 16S rRNA gene sequences that did not match sequences in the Greengene taxonomic database.

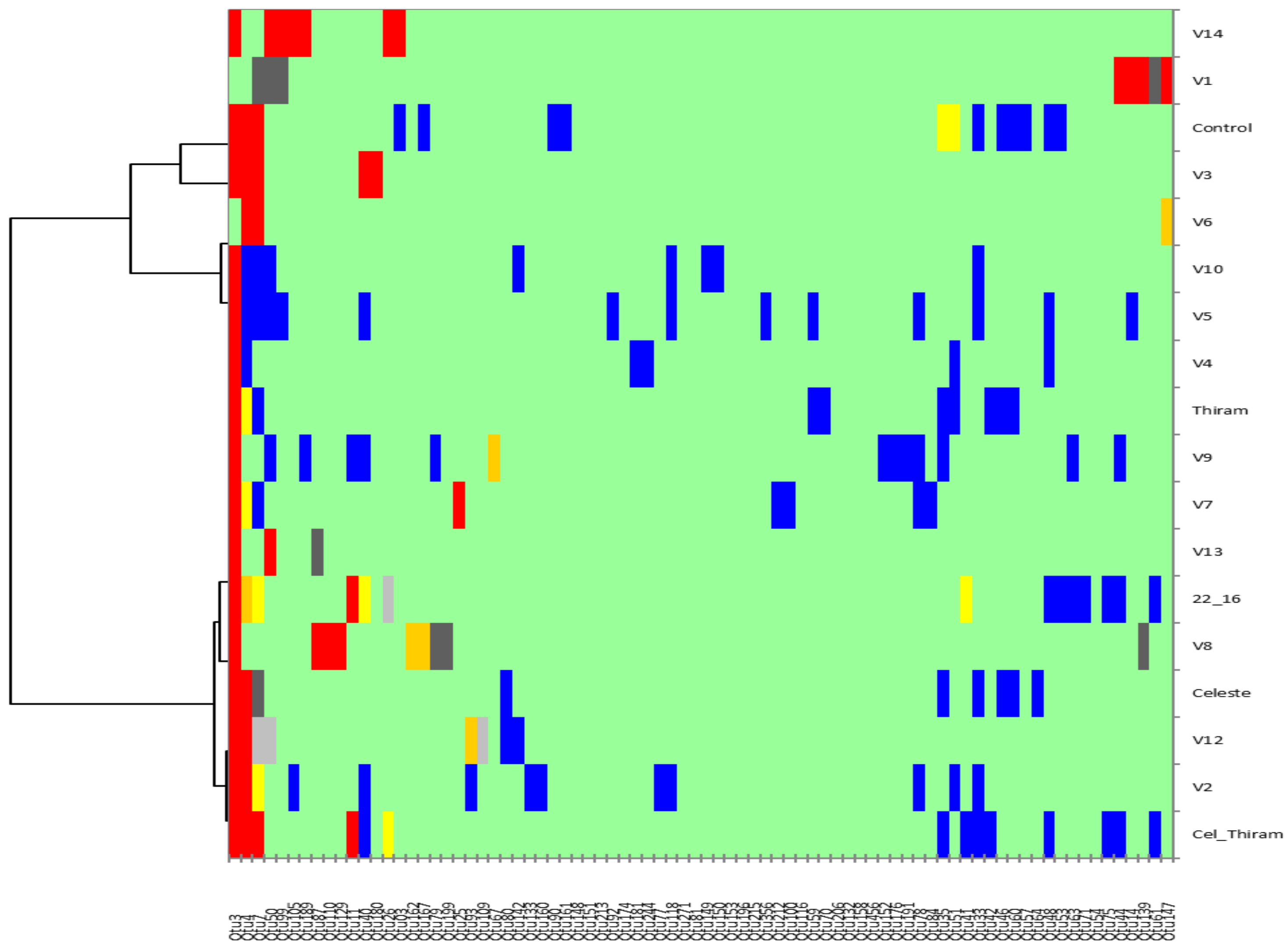

V14
V1
Control
V3
V6
V10
V5
V4
Thiram
V9
V7
V13
22_16
V8
Celeste
V12
V2
Cel_Thiram

Figure 3.2 Heatmap showing the diversity of selected bacterial operational taxonomic units (OTUs) detected in 18 onion seed lots of the same cultivar produced in two provinces in South Africa, using Illumina MiSeq analysis of the 16S rRNA sequences amplified from DNA extracted from each seed lot. The heatmap was constructed using xlstat (https://www.xlstat.com/). Seed lots are listed on the right vertical axis, and OTUs on the x-axis. Refer to Table 1 for details of the seed lots and Table S1 for identified OTUs.

3.9 Supplementary information

Table S1. Bacterial operational taxonomic units detected from 18 onion seed lots of the same proprietary cultivar produced in two provinces in South Africa.

OTU	Size[a]	Phylum	Family	Genus	Species
OTU0003	31209	Proteobacteria	Enterobacteriaceae (100)	*Pantoea* (61)	unclassified
OTU0004	11440	Proteobacteria	Pseudomonadaceae(100)	*Pseudomonas* (100)	unclassified
OTU0007	242	Proteobacteria	Xanthomonadaceae (100)	*Stenotrophomonas* (100)	*retroflexus* (66)
OTU0011	165	Bacteroidetes	Sphingobacteriaceae (100)	*Sphingobacterium* (100)	*multivorum* (100)
OTU0025	71	Bacteroidetes	Flavobacteriaceae (100)	*Myroides* (100)	unclassified (100)
OTU0026	71	Firmicutes	Bacillaceae (100)	*Bacillus* (98)	unclassified (98)
OTU0033	50	Firmicutes	Exiguobacteraceae (100)	*Exiguobacterium* (100)	unclassified (100)
OTU0035	40	Proteobacteria	Moraxellaceae (100)	*Acinetobacter* (100)	unclassified (68)
OTU0040	33	Proteobacteria	Oxalobacteraceae (100)	*Massilia* (97)	unclassified (88)
OTU0041	33	Bacteroidetes	Flavobacteriaceae (100)	*Flavobacterium* (100)	unclassified (100)
OTU0042	32	Proteobacteria	Rhodobacteraceae (100)	*Rhodobacter* (72)	unclassified (72)
OTU0044	31	Proteobacteria	Moraxellaceae (100)	*Acinetobacter* (100)	unclassified (71)
OTU0046	26	Proteobacteria	Enterobacteriaceae (100)	*Providencia* (93)	unclassified (93)
OTU0048	26	Firmicutes	Enterococcaceae (100)	*Enterococcus* (100)	*casseliflavus* (100)
OTU0050	24	Proteobacteria	Pseudomonadaceae (75)	*Pseudomonas* (75)	unclassified (100)
OTU0051	23	Bacteroidetes	eeksellaceae] (100)	*Wautersiella* (100)	unclassified (100)
OTU0053	22	Proteobacteria	Aeromonadaceae (100)	*Aeromonas* (100)	unclassified (100)
OTU0054	21	Proteobacteria	Nitrosomonadaceae (100)	*Nitrosomonas* (100)	*oligotropha* (100)
OTU0057	17	Proteobacteria	Sphingomonadaceae (100)	*Sphingomonas* (100)	*asaccharolytica* (60)
OTU0059	16	Proteobacteria	Halomonadaceae (100)	*Haererehalobacter* (100)	*salaria* (100)

OTU0060	15	Proteobacteria	Brucellaceae (100)	*Ochrobactrum* (100)	unclassified (100)
OTU0061	14	Proteobacteria	Caulobacteraceae (100)	*Brevundimonas* (100)	*diminuta* (100)
OTU0063	14	Proteobacteria	Bradyrhizobiaceae (100)	*Balneimonas* (100)	unclassified (100)
OTU0064	13	Actinobacteria	Nocardiopsaceae (100)	*Prauseria* (93)	unclassified (93)
OTU0067	12	Firmicutes	Leuconostocaceae (100)	*Weissella* (100)	unclassified (100)
OTU0070	11	Proteobacteria	Alcaligenaceae (100)	*Tetrathiobacter* (100)	*kashmirensis* (100)
OTU0071	10	Proteobacteria	Alcaligenaceae (100)	*Achromobacter* (100)	unclassified (100)
OTU0075	10	Proteobacteria	Burkholderiaceae (100)	*Burkholderia* (100)	unclassified (100)
OTU0077	9	Proteobacteria	Rhodobacteraceae (100)	*Paracoccus* (56)	unclassified (100)
OTU0078	9	Firmicutes	Veillonellaceae (100)	*Veillonella* (100)	*dispar* (100)
OTU0079	9	Firmicutes	Planococcaceae (100)	*Planomicrobium* (67)	unclassified (67)
OTU0080	9	Bacteroidetes	Weeksellaceae (100)	*Chryseobacterium* (100)	unclassified (100)
OTU0081	8	Proteobacteria	Sphingomonadaceae (100)	*Novosphingobium* (100)	unclassified(100)
OTU0084	8	Actinobacteria	Micrococcaceae (100)	*Microbispora (63)*	*rosea* (63)
OTU0087	7	Proteobacteria	Methylobacteriaceae(100)	*Magnetospirillum* (100)	*magnetotacticum* (100)
OTU0090	7	Bacteroidetes	Flavobacteriaceae(100)	*Flavobacterium* (100)	*succinicans* (100)
OTU0092	7	Firmicutes	Staphylococcaceae(100)	*Staphylococcus* (100)	unclassified (60)
OTU0093	7	Proteobacteria	Moraxellaceae (100)	*Psychrobacter* (100)	unclassified (100)
OTU0099	6	Cyanobacteria	Phormidiaceae (100)	*Planktothrix* (100)	unclassified (100)
OTU0100	6	Proteobacteria	Sphingomonadaceae (100)	*Sphingomonas* (100)	*wittichii* (100)
OTU0103	6	Bacteroidetes	Sphingobacteriaceae (100)	*Sphingobacterium* (100)	unclassified (100)
OTU0105	5	Proteobacteria	Moraxellaceae (100)	*Acinetobacter* (100)	*johnsonii* (100)
OTU0109	5	Proteobacteria	Pseudomonadaceae (80)	*Pseudomonas* (80)	unclassified (61)
OTU0110	5	Actinobacteria	Micrococcaceae (100)	*Arthrobacter* (80)	*agilis* (80)
OTU0114	5	Bacteroidetes	Sphingobacteriaceae (100)	*Sphingobacterium* (100)	unclassified (100)
OTU0116	5	Bacteroidetes	Prevotellaceae (100)	*Prevotella* (100)	unclassified (100)

OTU0118	5	Bacteroidetes	Prevotellaceae (100)	*Prevotella* (100)	*nanceiensis* (100)
OTU0129	4	Bacteroidetes	Bacteroidaceae (100)	*Bacteroides* (100)	unclassified (100)
OTU0132	4	Actinobacteria	Streptomycetaceae (100)	*Streptomyces* (100)	unclassified (100)
OTU0133	4	Firmicutes	Streptococcaceae (100)	*Streptococcus* (100)	unclassified (100)
OTU0139	3	Bacteroidetes	Weeksellaceae (100)	*Wautersiella* (100)	unclassified (100)
OTU0142	3	Proteobacteria	Methylobacteriaceae (100)	*Magnetospirillum* (100)	*magnetotacticum* (100)
OTU0147	3	Proteobacteria	Hyphomicrobiaceae (100)	*Hyphomicrobium* (100)	unclassified (100)
OTU0148	3	Proteobacteria	Sphingomonadaceae(100)	*Sphingobium* (100)	*yanoikuyae* (100)
OTU0150	3	Proteobacteria	Sphingomonadaceae (100)	*Sphingomonas* (100)	unclassified (100)
OTU0152	3	Actinobacteria	Brevibacteriaceae (100)	*Brevibacterium* (100)	*aureum* (100)
OTU0153	3	Actinobacteria	Cellulomonadaceae (100)	*Cellulomonas* (100)	*xylanilytica* (67)
OTU0158	3	Proteobacteria	Shewanellaceae (100)	*Shewanella* (100)	unclassified (100)
OTU0160	3	Actinobacteria	Promicromonosporaceae (67)	*Cellulosimicrobium* (67)	unclassified (67)
OTU0161	3	Bacteroidetes	Bacteroidaceae (100)	*Bacteroides* (100)	unclassified (100)
OTU0162	3	Bacteroidetes	Cytophagaceae (100)	*Hymenobacter* (100)	unclassified (100)
OTU0167	3	Proteobacteria	Brucellaceae (100)	*Ochrobactrum* (100)	intermedium(100)
OTU0174	3	Firmicutes	Paenibacillaceae (100)	*Paenibacillus* (100)	unclassified (100)
OTU0176	3	Actinobacteria	Bogoriellaceae (100)	*Georgenia* (100)	unclassified (100)
OTU0180	3	Firmicutes	Bacillaceae (100)	*Bacillus* (100)	*clausii* (100)
OTU0181	3	Firmicutes	Streptococcaceae (100)	*Streptococcus* (100)	*infantis* (67)
OTU0189	2	Proteobacteria	Comamonadaceae (100)	*Hydrogenophaga* (100)	unclassified (100)
OTU0191	2	Bacteroidetes	Sphingobacteriaceae (100)	*Sphingobacterium* (100)	*mizutaii* (100)
OTU0196	2	Actinobacteria	Streptomycetaceae (100)	*Streptomyces* (100)	unclassified (100)
OTU0199	2	Proteobacteria	Methylophilaceae (100)	*Methylobacillus* (100)	unclassified (100)
OTU0206	2	Proteobacteria	Sphingomonadaceae (100)	*Sphingomonas* (100)	unclassified (100)
OTU0212	2	Proteobacteria	Xanthomonadaceae (100)	*Stenotrophomonas* (100)	*retroflexus* (100)

OTU0213	2	Firmicutes	Lactobacillaceae (100)	*Lactobacillus* (100)	unclassified (100)
OTU0215	2	Planctomycetes	Planctomycetaceae(100)	*Planctomyces* (100)	unclassified (100)
OTU0271	2	Proteobacteria	Pseudomonadaceae (100)	*Pseudomonas* (100)	*stutzeri* (100)
OTU0244	2	Proteobacteria	Pseudomonadaceae (100)	*Pseudomonas* (100)	unclassified (100)
OTU0356	1	Proteobacteria	Pseudomonadaceae (100)	*Pseudomonas* (100)	*veronii* (100)
OTU0371	1	Proteobacteria	Pseudomonadaceae (100)	*Pseudomonas* (100)	unclassified (100)
OTU0456	1	Proteobacteria	Xanthomonadaceae (100)	*Xanthomonas* (100)	*campestris* (100)

[a] Size = how many times an OTU was observed in the 18 seed lots, e.g., OTU0003 was observed 31209 times and it is 61% identical to *Pantoea* spp.

Chapter 4

Comparative genomics of pathogenic and non-pathogenic *Pantoea agglomerans* strains isolated from onion seeds

4.1 Abstract

Pantoea agglomerans is an ubiquitous bacterium commonly isolated from plants, seeds, fruits and even humans. The majority of strains are non-pathogenic epiphytes and endophytes. Some are beneficial and have been reported as plant growth promoters and/or biological control agents. In onion, some strains of *P. agglomerans* cause leaf blight and seed stalk necrosis. Both pathogenic and non-pathogenic strains can exist in the same environment, e.g., in onion seed. In this study, the whole genomes of two strains of *P. agglomerans*, one non-pathogenic on onion (BD 1212) and one pathogenic on onion (BD 1274) that had been isolated from the same onion seed lot were compared. Genes that differed between the pathogenic and non-pathogenic strains were identified. Genes coding for non-fimbrial adhesins and a non-functional (incomplete) Type III secretion system were present in the non-pathogenic strain BD 1212, but not in the pathogenic strain BD1274. The pathogenic strain contained genes involved in conjugal transfer (Type IV secretion system) and production of four toxins, which were absent in the non-pathogenic strain. The genes unique to each strain may contribute to adaptation mechanisms and fitness of the non-pathogenic strain, and understanding virulence factors of the pathogenic strain to onion plants.

4.2 Introduction

Pantoea agglomerans (formerly *Erwinia herbicola*) is a member of the Enterobacteriaceae. This bacterium has been isolated as a symbiont from wheat (*Triticum*) (Lindh et al., 1991), rice (Feng et al., 2006) and sweet potato (*Ipomoea batatas*) (Asis & Adachi, 2004) and some strains have been shown to have plant growth promoting activity. *Pantoea agglomerans* also causes diseases in cotton (*Gossypium* L.) (Medrano et al., 2007), maize (*Zea mays*) (Morales-Valenzuela et al., 2007) and rice (*Oryza sativa*) (González et al., 2015). In onions, some strains of *P. agglomerans* cause leaf blight and seed stalk necrosis (Edens et al., 2006; Hattingh & Walcott, 1981). Some strains of this bacterium are also used as biological control agents of postharvest diseases of pome fruit, while others cause human diseases (Cruz et al., 2007; Nunes et al., 2002).

Bacteria colonising seeds have been reported since the 1970s (Mundt & Hinkle, 1976), however only few studies have been conducted on bacteria occurring in seeds as endophytes, and their origin in this environment is being debated (Barret et al., 2015; Cesbron et al., 2015;

Hardoim et al., 2012; Johnston-Monje & Raizada, 2011; Załuga et al., 2014). Understanding the characteristics/features and actions of seedborne microorganisms is necessary for the determination of their successful transmission to the next generation of plants (Barret et al., 2016).

Several research groups have explored factors associated with the transmission of bacteria to seed and from seed to plants. *Xanthomonas* spp. and *Acidovorax* spp. effectively colonise a germinating host seedling using quorum sensing and secretion of microbial effectors and adhesion to a cell (Barret et al., 2016; Darsonval et al., 2008, 2009; Johnson & Walcott, 2013). Barret et al. (2016) and Truyens et al. (2015) reported that chemotaxis, metabolism of carbohydrates and iron uptake are vital for the successful colonisation of germinating seeds by endophytes.

Moretti et al. (2014b) sequenced the full genome of a *P. agglomerans* strain DAPP-PG734 isolated from an olive plant affected by the knot disease. The strain contained a complete *hrc/hrp* gene cluster, similar to those present in *Erwinia amylovora* and *E. pyrifoliae* causing fire blight of pears (Moretti et al., 2014a). Pathogenicity of *P. agglomerans* pv. *gypsophilae* 824-1 and *P. agglomerans* pv. *betae* 4188 causing galls in gypsophila and beet respectively, is due to the presence of plasmid pPATH (Manulis & Barash, 2003). Manulis & Barash (2003) and Barash & Manulis-Sasson (2009) reported that the pathogenicity islands present in this plasmid harbour the *hrp* gene cluster, which are involved in encoding type III effector proteins.

In this study, we present the whole genome analysis of two strains of *P. agglomerans*, non-pathogenic BD 1212 and pathogenic BD 1274 causing leaf and seed stalk blight on onion, isolated from the same seed lot. The aim of this work was to identify genomic differences between pathogenic and non-pathogenic strains in order to reveal possible genetic factors important for emergence of pathogenicity on onion.

4.3 Materials and methods

Bacterial strains

Pantoea agglomerans strains BD 1212 (non-pathogenic on onion) and BD 1274 (pathogenic on onion), used in this study were isolated from the same onion seed lot obtained from a seed producer in the Western Cape, South Africa. The bacterial strains were stored in the Plant Pathogenic and Plant Protecting Bacteria (PPPPB) National Collection at the Agricultural Research Council: Plant Health and Protection in Pretoria, South Africa. In a pathogenicity trial as described by Goszczynska et al. (2006) and Moloto et al. (2017), BD 1212 did not induce symptoms when inoculated onto onion seedlings, whereas BD 1274 was pathogenic. The strain BD 1274 induced water soaked lesion onto onion seedlings which after six days turned to necrotic (As described in Chapter 2).

Genomic DNA extraction, sequencing, genome alignment and annotation

For genomic DNA extraction, the strains were streaked on TGA (Difco) medium; plates were incubated overnight at 28°C. Genomic DNA from strains BD 1212 and BD 1274 was extracted from a single colony of each strain using the Wizard® Genomic DNA purification kit (Promega, Madison, WI), according to the manufacturer's instructions. The quality and quantity of the genomic DNA were evaluated on agarose gels and using a Qubit® 2.0 fluorometer (Life Technologies, Carlsbad, US-CA), respectively. The genomes of the two strains were sequenced using mate-paired Illumina sequencing using the HiSeq 2500 platform following Inqaba Biotechnology (South Africa) protocol. Libraries with an insert size of 500 bp were generated and sequence lengths of 90 bp in both directions were obtained.

The pair end reads were assembled into contigs using Spades v3.9.0 (http://bionif.sp.bau.ru/en/spades) and the read quality was determined using Trimmomatic v.0.39 (Bolger et al., 2014). Assembled genomes were aligned using multiple alignment of conserved genomic sequence (Mauve) version 2.1.0 (Darling et al., 2004). Briefly, Mauve was opened, align with progressive Mauve was selected. A reference genome sequence and annotated genomes were downloaded in the program to be aligned. A name for the output file was written before it was aligned. When alignment is completed, the alignment file open automatically. Each row represent a genome and each coloured block represent genetic similarity. Coloured blocks in the first genome are connected by lines to similarly colored

blocks in the second and third genomes (Darling et al., 2004). These lines indicate which regions in each genome are similar.

The genomes were annotated using the RAST v2.0 (Rapid Annotation using Subsystem Technology, http://rast.nmpdr.org/), server (Aziz et al., 2008) and PATRIC v3.6.3 (Pathosystem Resource Integration Centre) annotation (www.patricbrc.org Wattam et al., 2014). To annotated the genomes with RAST server, users are directed to the "Jobs Overview once logged onto the server. A new job was started, the genomes to be uploaded were selected from the navigation bar or the link near the top of the page. Valid taxonomy id, the organism's genus, species, and strain, as well as a nucleotide sequence file in FASTA format were provided. After the annotation is complete, the annotated genome were download in a variety of export formats (e.g. GenBank, FASTA, Excel) or browse the genome in the comparative environment of the SEED-Viewer

The genome annotation service in PATRIC (Wattam et al., 2014) uses the RAST tool kit (RASTtk) [Brettin et al., 2015] to provide annotation of genomic features. Briefly, the genomes were uploaded into a private workspace of the annotation service. Up load the contig file (Fasta format) by clicking on the arrow at the end of the text box with the word Unspecified in it. Specific name that will identify the genome was entered in the workplace, once the name is entered; the OUTPUT NAME shows the selected genus, species and identifier. The default genetic code is 11, for Archaea and most Bacteria was selected. An output folder where the annotation will be placed was allocated. The annotation job was started by clicking on the annotate button. Once the annotation was completed, a file feature.txt was downloaded. The file is a tab-delimited text file listing all the features of the genome. Each feature contains the PATRIC ID, the location string, the feature type, the functional assignment, any alternated IDs found, and protein-coding genes.

The organism overview page containing basic information on the genome such as taxonomy, size, the number of contigs, the number of coding sequences and RNAs and counts of non-hypothetical and hypothetical gene annotations was presented in Table 4.1. These features obtained from both RAST and PATRIC annotations were compared manually, all the proteins genes including hypothetical genes were blasted using BLASTn and BLASTp. Conserved proteins were confirmed using NCBI conserved domain search (www.ncbi.nlm.nih.gov), and the genes that were unique between the two genomes were confirmed.

Genome accession numbers

The Whole Genome Shotgun sequencing project were deposited at DDBJ/EMBL/GenBank under the accessions QQXH00000000 (BD 1212) and QQXI00000000 (BD 1274).

Genomic comparisons

To visualize differential gene content between the two genomes we uploaded in EDGAR (Efficient Database framework for comparative Genome Analyses using BLAST score Ratios) (Blom et al., 2009) and create Venn diagrams for the two genomes. The Venn diagrams show the numbers of reciprocal best hits between subsets of genomes. Every area in this Venn diagram represents a subset of the compared genomes and is labeled with the number of genes in this subset. To simplify the assignment of an area to a genome set every genome has a base color. The areas of the Venn diagram are colored in the averaged color of the associated genomes. On the left side a Venn diagram of two *Pantoea agglomerans* genomes is shown that illustrates the number of singleton genes (pink and blue) and the core genome (centered yellow). Unique protein coding sequences (CDSs) predicted by PATRIC analysis (Wattam et al., 2014) were confirmed manually by BLASTp (blast.ncbi.nlm.nih.gov). Proteins with amino acid similarities greater than 50% and with a coverage of 70% were considered homologs.

A circular genome map was constructed using the BLAST Ring Image Generator (BRIG v0.95) following protocol (Alikhan et al., 2011; https://sourceforge.net/projects/brig/files/). The circular genome was drawn using the complete genome of *P. agglomerans* strain C410P1 (NZ_CP016889.1) (Li, 2016; unpublished) as a reference genome on BLAST. Briefly, download BRIG-Example.zip file, which is available from the BLAST Ring Image Generator (BRIG) website: https://sourceforge.net/projects/brig/files/). Unzip it where is easily accessible, like the home directory or desktop. The sequence were loaded, BRIGExample.fna was fixed as the reference sequence. Add to query pool was pressed, this loaded several items into the pool list. In the next step, Information shown on each of the concentric rings in BRIG was arranged. Three rings were created as follow, for each ring: the legend text was adjusted for each ring. The required sequences was selected from the data pool, then add data to add to the ring list was pressed. A colour of choice was chosen. The upper (90) and lower (70) identity threshold were adjusted. Click on “add new ring” and repeat steps for each new ring required. Each of the rings were arranged as follows: Ring 1- Legend text: GC Content Required sequences: GC Content Ring 2: Legend text: GC Skew Required sequences: GC Skew. The image title was positioned as “BRIG example image”. Submit button was clicked, then BRIG

formated Genbank files, performed BLAST, parsed the results and rendered the image (Alikhan et al. 2011; http://brig.sourceforge.net/). The image was created in the specified output directory and is shown in Figure 4.4.

4.4 Results

General features of the genome sequences

The *P. agglomerans* BD 1212 genome was assembled into 103 contigs with a total of 4 875 404 bp and G+C content of 55.1%, while that of BD 1274 was assembled into 246 contigs, with a genome of 4 968 508 bp and a mean G+C content of 55% (Table 4.1). Of the 4751 genes detected in the genome of BD 1212, 4492 genes (94.54%) were protein-coding sequences, of which 80.17% had a predicted function (Table 1). Of 4918 genes detected in the genome of BD 1274, 4568 genes (92.88%) were protein coding sequences, of which 84.01% had a functional prediction (Table 4.1). Fifty-six RNA genes were identified in the genome of BD 1212, including four ribosomal RNA (rRNA) and 52 transfer RNA (tRNA) genes; while 64 RNA genes were identified in the genome of BD 1274, including 11 rRNA and 53 tRNA genes (Table 4.1). The ribosomal RNA operon showed a characteristic bacterial organisation with genes for 5S, 16S and 23S rRNA and the tRNAs tRNAIle and tRNAAla present in both of the genomes.

According to the Venn diagram (Fig 4.1), the two strains shared 4002 protein coding genes, corresponding to 87 to 89% of all CDS in these genomes. Seventy-nine unique genes were present only in the non-pathogenic strain and were involved in carbohydrate metabolism, cell wall synthesis, Type III secretion system (T3SS), cell adhesion, bacteriophages/phage/mobile genetic elements and other genes (Table 4.2). Sixty unique genes were identified only in the pathogenic strain and are involved in carbohydrate metabolism, Type IV secretion system (T4SS), toxins production, mobile genetic elements, vitamins and other genes responsible for virulence, excluding hypothetical or unknown proteins (Table 4.3). BLASTn revealed that BD 1212 and BD 1274 genomes each had three plasmids with a sequence similarity of 98% to CP016890.1, CP016891.1 and CP016892.1 plasmids in *P. agglomerans* C410P1. Additionally, the pathogenic strain BD 1274 had a fourth plasmid, CP014126.1, similar to that in *P. agglomerans* strain FDAARGOS-160. In addition, the genome scale comparison of draft genome sequences of *P. agglomerans* BD 1212 and BD 1274 showed that the strains differed

by the presence of unique segments in the pathogenic strain (Fig 4.2). Comparative circular BLAST alignment of the genome structure (Fig 4.3) showed high homology between BD 1212 and BD 1274 compared to the complete reference genome of *P. agglomerans* C410P1 (NZ_CP016889.1).

Comparative analysis

Carbohydrates play several crucial roles in the metabolic processes of living organisms, serving as energy sources (Palmer et al., 2018). Comparison of functional categories between the genomes of the two onion seed strains of *P. agglomerans* showed that the largest number of genes in the two genomes were involved in carbohydrate metabolism (Fig 4.4; Table S1), such as the phosphotransferase system, 3-keto-L-gulonate/L-ascorbate specific IIA component (PtxA), sucrose-6-phosphate hydrolase, GDP-mannose 4,6-dehydratase, trehalose-6-phosphate phosphatase, D-ribose pyranase, glucosyl-3-phosphoglycerate synthase and myo-inositol 2-dehydrogenase. A difference was seen in the mannose pathway as mannose-1-phosphate guanylyltransferase was identified only in the non-pathogenic strain BD 1212 (Table 2). Some proteins related to the pentose phosphate pathway, viz. ribose 5-phosphate isomerase B, sucrose metabolism (sucrose permease), glycolysis and gluconeogenesis metabolism (unsaturated glucuronyl hydrolase and xylan 1,4-beta-xylosidase), glyoxylate dicarboxylaste metabolism (putative N-acetylglucosamine kinase) and nucleotide metabolism (Deoxythymidine diphosphate (dTDP)-4-dehydrorhamnose 3,5-epimerase) were identified in the pathogenic strain BD 1274 only (Table 4.3).

Insertion sequence elements (IS)

Insertion sequences (IS) are fragments of other mobile genetic elements such as phages and plasmids and were reported to be transferred horizontally between genomes (Siguier et al. 2014; Vandecraen et al., 2017). Insertion sequence element/transposase related proteins such as transposase STM474_p1058, the IS3 transposase family, IS3/IS911 family, and ISNCY family were present in both strains (Table S2). In comparison, the transposase InsK of IS150, InsO and ISL3 were present only in the pathogenic strain (Table 3). A BLASTp search showed that ISL3 belongs to an enterotoxin. Enterotoxins are exotoxins produced and secreted by several bacterial species (Vandecraen et al., 2017). IS1 related to InsA, ISSod13, transposase InsF for insertion sequence IS3, transposase ISL3, and transposase IS66 were present only in the non-pathogenic strain (Table 4.2).

Plant cell wall degrading enzymes

Plant cell wall degrading enzymes release nutrients and water from plant cells to ensure effective colonisation of the tissues by bacteria (Bender, 1999; Melotto & Kunkel, 2013). Genes encoding cell wall degrading enzymes such as amylase (glucoamylase), cytoplasmic alpha-amylase, cellulase (beta-1,4-glucanase), xylan (endo-1,4-beta-xylanase), pectin (polygalacturonase, a putative pectin degradation protein), proteases (metalloprotease, exported zinc metalloprotease YfgC precursor, HtrA protease/chaperone protein, and ClpXP protease specificity-enhancing factor/stringent starvation protein B) were identified in both genomes (Table S2). The difference between the two genomes in terms of cell wall degrading enzymes was a protein encoding oligogalacturonate-specific porin identified only in the non-pathogenic strain (Table 4.2).

T3SS

Many Gram negative bacteria use T3SSs to transfer effectors (T3Es) proteins straight into the cytoplasm of infected host cells and colonise the host cells (Alfano & Collmer, 2004). Candidate type III effector Hop (hrp-dependent outer protein) protein was present in both genomes (Table S2). However, an incomplete T3SS (Shariati et al., 2017) was present only in the non-pathogenic strain BD 1212 and contained four *hrp* genes (*hrpB*, *hrpD*, *hrpF*, *hrpJ*), the *hrc* gene (*hrcQb*), the type III secretion spans bacterial envelope protein (YscO), the type III secretion bridge between inner and outer membrane lipoprotein (YscJ,HrcJ,EscJ,PscJ) and the type III secretion inner membrane protein (Table 4.2). There was no evidence of a T3SS in the genome of the pathogenic strain BD 1274. Additionally, five proteins of the type III export apparatus that are homologous to the flagellar component were present in the non-pathogenic strain. They included the type III secretion inner membrane protein (YscR,SpaR, HrcR,EscR, homologous to the flagellar biosynthesis protein FliP), type III secretion inner membrane protein (YscS, homologous to the flagellar biosynthesis protein FliQ), type III secretion inner membrane protein (YscT,HrcT,SpaR,EscT,EpaR1, homologous to the flagellar biosynthesis protein FliR), the type III secretion inner membrane protein (YscU,SpaS,EscU, HrcU,SsaU, homologous to the flagellar biosynthesis protein FlhB) and type III secretion outer membrane pore forming protein (YscC,MxiD,HrcC,InvG), homologous to the type IV pilus biogenesis protein PilQ (Table 4.2).

T4SS

Type IV secretion systems (T4SSs) play an essential role in pathogenesis (Christie 2016; Christie et al., 2005; Lawley et al., 2003). In this study, Type IV secretion transport related proteins homologous to the conjugative transfer of pilus assembly (TraC and TraV) were found in both the BD 1212 and BD 1274 genomes (Table S2). The difference between the genomes was a set of proteins involved in conjugal transfer that was identified only in the pathogenic strain BD 1274 (Table 4.3). These included the conjugative transfer of plasmid DNA (TraN, -G, -D), transfer of pilus assembly (TraB, -F, -C, -H and -W), surface exclusion (TraT), and DNA-nicking and unwinding (TraI) proteins. The type-F conjugative transfer system pilin assembly protein/thiol-disulfide isomerase, TrbC and TrbB were also identified in the pathogenic strain (Table 4.3).

Flagellar structure

Motility allows bacteria to spread swiftly between and within host plants (Diepold & Armitage, 2015; Weller-Stuart et al., 2017). Furthermore, flagella play a role in attachment, biofilm formation as well as swimming motility (Diepold & Armitage, 2015; Haiko et al., 2013). Proteins related to flagellar assembly in the core genomes of *P. agglomerans* consist of the flagellar biosynthesis proteins FlhAQRO; flagella basal body P-ring formation proteins FlgAC; flagellar assembly protein FliH; flagellar basal body rod protein components FlaE and FlgJ; flagellarhook-basal body complex proteins FliELK and FlgCK; flagellar related proteins FliJ and FlhE; flagellar motor proteins MotA/MotB; and FliNGMSTZ which were identified in both genomes. The flagellin related proteins FlaA and FliC and a negative regulator of flagellin synthesis FlgM (anti-sigma28) were also present in both genomes (Table S2). A protein related to twitching motility, PilT, was also present in both genomes (Table S2).

Type IV pilus biogenesis proteins, fimbriae and non-fimbrial adhesins

Bacterial attachment is an essential first step for most bacteria to colonise and persist within the host (Pizarro-Cerdá, 2006; Vo et al., 2017). Non-fimbrial (autotransporters, and filamentous haemagglutinin-like) or fimbrial (incuding type IV pili) adhesin-related proteins such as the type IV pilin, PilA, type IV pilus biogenesis proteins such as PilMQ, type 1 fimbriae major subunit FimACD, type IV fimbrial assembly protein/ATPase *PilCB,* the sigma-fimbriae usher protein, the sigma-fimbriae tip adhesion and putative pilus chaperone, and the PapD family were found in both non-pathogenic and pathogenic strains (Table S2). Only the non-pathogenic strain BD 1212 had putative fimbrial structural unit proteins that are homologs of

AfaD, as well as a filamentous haemagglutinin family outer membrane protein encoding EndoU nuclease, the adhesin autotransporter YdeU, AidA-1 and an uncharacterized protein YcgV that is homologous to AiDA (Table 4.2).

Toxin-antitoxin

Most bacterial genomes have multiple type II toxin–antitoxin systems (TAs), encoding two proteins named as a toxin and an antitoxin (Pandey & Gerdes, 2007; Ramisetty & Santhosh, 2017; Van Melderen & de Bast, 2009; Yamaguchi et al., 2011). Toxins inhibit cellular processes of the host, while the relation of the antitoxin with the toxin reduces the toxin's activity (Ramisetty & Santhosh, 2017). In this study, toxin-related proteins such as Orphan toxin OrtT and other toxin-antitoxin related proteins, such as CcdA protein (antitoxin to CcdB), toxin HigB (antitoxin HihA), vapC toxin, ribosome association toxin RatA, per-activated serine protease autotransporter enterotoxin EspC and RelE-like translational repressor toxin, were present in both the onion non-pathogenic and onion pathogenic strains of *P. agglomerans* (Table S2). The only differences were the toxin probable mRNA interferase HicA, phage DNA binding protein HicB, YefM protein (antitoxin to YoeB) and EF hand domain protein (pesticin, a bacterial toxin homolog of phage lysozyme), all of which were present only in the genome of the pathogenic strain BD 1274 (Table 4.3).

Other mobile genetic elements associated with virulence on onion

Bacteriophages/phages/integrase play a major role in bacterial development through the horizontal transfer of genes that contribute to fitness and pathogenesis of the bacterial pathogen (Buttimer et al., 2017; Christie & Calendar, 2016; Evans et al., 2010; Griffith et al., 2000; Nilsson & Haggård-Ljungquist, 2007). Forty-seven bacteriophage/phage related proteins were detected in both strains BD 1212 and 1274 of *P. agglomerans* (Tables 4.2, 4.3, and S2). Thirty of these proteins were present only in the non-pathogenic strain BD 1212 (Table 4.2). They comprised the immunity repressor of bacteriophage P2, bacteriophage CI repressor, phage baseplate assembly protein V/Gp45, and gp3 phage tail completion protein that probably contribute to the fitness of the non-pathogenic bacterium in different environments (Buttimer et al., 2017; Christie & Calendar, 2016; Evans et al., 2010; Nilsson & Haggård-Ljungquist, 2007). Grainge & Sherratt (1999) reported that Lambdoid phage Rac integrase is involved in cleavage of a single strand of a DNA duplex by nucleophilic attack of a sealed tyrosine to give a 3' phosphotyrosyl protein-DNA. This protein was identified only in the pathogenic strain BD 1274 (Table 4.3).

4.5 Discussion

Sequencing the whole genome facilitated, comparative genome analysis of increasing number of genomes, whereby comparison of one pathogen to the genome of other sequenced pathogens can reveal important virulent factors shared by the two organism. In this study, we sequenced and performed comparative analysis of pathogenic (BD 1274) and non-pathogenic *P. agglomerans* strain in order to identify different genes between the genomes, Genomic analysis revealed that strain BD 1274 of *P. agglomerans*, which is pathogenic on onion, has a larger genome (4, 968 508 bp) than the non-pathogenic strain BD 1212 (4, 875 404 bp), confirming prior observations that non-pathogens have reduced genomes compared to pathogenic strains of some bacteria (Cesbron et al., 2015; Lòpez-Fernandez et al., 2015). The fourth plasmid in the pathogenic strain contains a cluster of genes which are responsible for the conjugal transfer of DNA, playing major role in pathogenicity. The possible explaination for the non-pathogenic nature of strain BD 1212 is that it didn't induce symptoms when inoculated into onion and spring onion seedling and the lack of the fourth plasmid carrying the T4SS genes responsible for pathogenicity or the four toxin production proteins. Despite the overall high degree of similarity in virulence related genes detected in the pathogenic and non-pathogenic strains, genes unique to each strain were identified. The unique genes present in the pathogenic and non-pathogenic strains, may be important in general bacterial fitness, colonisation, survival and virulence in onion plants.

Comparison of the genomes of the two *P. agglomerans* strains isolated from the same onion seed lot showed the presence of genes encoding carbohydrate metabolism such as genes in the pentose phosphate pathway, sucrose metabolism, glycolysis and gluconeogenesis, glyoxylate and dicarboxylate metabolism and nucleotide pathways. These genes could play a major role in adaptation and survival of *P. agglomerans* in different environments. For example, deoxythymidine diphosphate (dTDP)-L-rhamnose is a sugar required for virulence in some pathogenic bacteria (Christendat et al., 2000; Rahim et al., 2001). A putative epimerase that converts N-acetylmannosamine-6-phosphate to N-acetylglucosamine-6-phosphate in the N-acetylmannosamine utilisation pathway was found in the pathogenic bacterium and allows the bacterial cells to use sialic acid as a carbohydrate source (Plumbridge & Vimr, 1999). A recent study showed that *Pantoea* species use carbohydrate derivatives such as oxaloacetate sucrose, which might enhance the ability of these bacteria to perform important cellular functions when nutrients are limited (Palmer et al., 2018).

Plant cell wall enzymes such as cellulases, pectinases, proteases and xylanases break down the plant cell wall and provide nutrients for the bacteria supporting the spread of the bacterium through the plant tissue (Korotkov et al., 2012; Pfeilmeier et al., 2016; Toth et al., 2003). Doughari (2015) and Melotto & Kunkell (2013) reported that, these enzymes play an important role in plant pathogen-interactions. These enzymes work together to break down plant cell walls, enabling the pathogen entry into host cells and the release of nutrients in plant cells (Cesbron et al., 2015; Melotto & Kunkell, 2013). For example, *Dickeya dadantii* feeds on plant cell walls by secreting pectinases and utilising the oligogalacturanate products (Blot et al., 2002; Hutter et al., 2014). Hutter et al. (2014) reported that *D. dadantii* produces a porin, enzyme degrading an outer membrane of the host cell enabling the uptake of oligosaccharides. In a study conducted by Cesbron et al. (2015), pectinase activity was only observed in non-pathogenic strains and it was speculated that this might play a role in nutrient uptake by these bacteria in planta. Based on the findings from this study with onion strains of *P. agglomerans*, the non-pathogenic *P. agglomerans* strain might use oligogalacturonate-specific porin (Table 4.2) for nutrient uptake and spread in onion cells (Cesbron et al., 2015; Vorhölter et al., 2012).

The T3SS transports virulence factors or effector proteins into host cells, triggering the host defence responses (Alfano & Collmer, 1997). For example, the T3SS is encoded by a *hrp* (HR and pathogenicity)/*hrc* (Hrp conserved) gene cluster in *Pseudomonas syringae* (Alfano & Collmer, 2004; Bonas & Lahaye, 2002). The *Hrc* proteins are conserved in all bacteria that possess a T3SS and the majority have homologous proteins that make-up the basal body of the flagellar biogenesis system (Büttner, 2012; Diepold & Armitage, 2015). An incomplete T3SS was found in the non-pathogenic onion strain of *P. agglomerans* in this study. Bacteria with an incomplete T3SS are unable to induce a plant hypersensitive response and basal defences (Alfano & Collmer, 2004; Benali et al., 2014; Melotto & Kunkell, 2013; Shariati et al., 2017).

Tampakaki (2014) reported that even though the role of the T3SS in non-pathogenic bacteria is still unknown, their existence may contribute to the environmental fitness of bacteria. For example, Marguerettaz et al. (2011) reported that the Inv/Mxi/Spa type T3SS present in *X. albilineans* is not essential for xylem colonisation and development of a leaf scald in sugarcane. The authors inoculated sugarcane with knockout mutants. The plants did not show any symptoms of the disease, which showed that the T3SS is not needed by *X. albilineans* to induce leaf scald in sugarcane (Marguerettaz et al., 2011). They also isolated strains of *X. albilineans* from diseased sugarcane that were lacking the T3SS SPI-1 (Marguerettaz et al., 2011). In other

studies, it was shown that the incomplete T3SS in some *Pseudomonas* strains was involved in biocontrol activity (Brader et al., 2017; Preston et al., 2001; Rezzonico et al., 2005; Shariati et al., 2017).

Shariati et al. (2017) compared strains of *P. agglomerans* isolated from different hosts, the authors identified an incomplete T3SS from a plant growth-promoting rhizobacterium strain P5 of *P. agglomerans* and two non-pathogenic strains of *P. agglomerans*, 190 and IG1. Thus, the presence of an incomplete T3SS in the onion non-pathogenic strain of *P. agglomerans* BD 1212 in this study may contribute to general fitness of the bacterium by increasing efficiency and adaptive abilities rather than play a role in virulence on onion. Several studies have shown that motility is a key feature influencing bacterial fitness and survival in different environments (Diepold & Armitage, 2015; Weller-Stuart et al., 2017). The presence of proteins related to the flagellar type III export apparatus in the onion strains of *P. agglomerans* evaluated in this study, such as the soluble proteins FliH and FliI, and the integral membrane proteins FlhA, FlhB, FliP, and FliQ in the non-pathogenic strain might play a significant role in swimming motility (Imada et al., 2015), biofilm formation and attachment to host cells (Diepold & Armitage, 2015; Haiko et al., 2013).

The T4SS functions in conjugal transfer of DNA between bacteria, but also in the transport of DNA and proteins from the bacterial cytoplasm to the host cell (Christie & Cascales, 2005; Lawley et al., 2003). The T4SS contributes to genome plasticity in bacteria and was reported to play a major role in pathogenesis in both plants and mammalian bacterial pathogens (Christie & Cascales, 2005; Christie et al., 2005; Liu et al., 2013). In this study, a set of proteins involved in conjugal transfer of DNA between bacteria were identified in the onion pathogenic strain of *P. agglomerans*, BD 1274. These proteins are important for pilus assembly and mating pair stabilisation (Lawley et al., 2003; Zechner et al., 2012). For example, TraC is required for the assembly of F pilin (Lawley et al., 2003). TraN is a cysteine-rich outer membrane protein involved in the mating-pair stabilisation (adhesin) component of the F-type conjugative plasmid transfer system (Lawley et al., 2003). TraD performs a coupling function, and is located at the inner membrane as well as the relaxosome-plasmid DNA complex (through TraM) (Beranek et al., 2004; Lawley et al., 2003). The TraG protein is believed to act together with the periplasmic domain of TraN to stabilize the mating-cell (Lawley et al., 2003).

The presence of gene clusters in the genome of pathogenic strain BD 1274 that look like both a plasmid conjugation system in *E. coli* (Lawley et al., 2003) and the pathogenicity-related T4SS locus in *A. tumefaciens* (Christie, 2016) suggested a potential role in pathogenicity to onion and horizontal gene transfer. Conjugation systems are responsible for horizontal gene transfer of antibiotic resistance genes, virulence factors and genes encoding other factors that benefit the bacterial cell (Christie, 2016; Grohmann et al., 2018; Souza et al., 2015). In *Xanthomonas citri*, the T4SS is used for toxin translocation to kill other bacteria and provide a competitive growth benefit in various bacterial communities (Grohmann et al., 2018; Souza et al., 2015). Further investigations are needed to clarify the exact functions of these genes in the onion pathogenic strain of *P. agglemerans*.

Bacterial attachment to plant surfaces is the initial step in biofilm formation and cell adhesion in several bacterial species (Berne et al., 2015; Sibanda et al., 2018; Vo et al., 2017). The filamentous hemagglutinin (FHA) proteins encoding the EndoU nuclease domain present in the non-pathogenic onion strain BD 1212 of *P. agglomerans* in this study are widely used by bacteria as toxins for defence, offense or addiction of selfish elements (Renzi et al., 2006; Zhang et al., 2011). The AidA-1 protein was reported to be involved in diffuse adherence in epithelial cells, aggregation, biofilm formation and pathogenicity in *E. coli* (Grijpstra et al., 2013; Henderson et al., 2004; Vo et al., 2017). The presence of AidA-1 in strain BD 1212 suggests this protein might play an essential part in colonisation and attack of host cells, along with allowing persistence of the non-pathogenic strain on onion seeds through biofilm formation and bacterial aggregates.

The Type II toxin-antitoxin (TA) system of the HicA family (Table 3) is part of the HicA-HicB TA systems that plays a role in inducing cleavage of mRNA and tmRNA (transfer-mRNA), thereby preventing translation to proteins (Jørgensen et al., 2009; Makarova et al., 2006; Ramisetty & Sonthosh, 2017). The hicAB locus is transcribed in response to amino acid and carbon starvation (Jørgensen et al., 2009; Makarova et al., 2006). Toxin-antitoxin (TA) system play a significant roles in bacterial immunity (Makarova et al., 2006; Ramisetty & Sonthosh, 2017; Van Melderen & de Bast, 2009; Yamaguchi et al., 2011). The presence of the orphan toxin OrtT (orphan toxin related to tetrahydrofolate) in the onion non-pathogenic strain of *P. agglomerans* suggested that this bacterium may use the TA system for survival during harvest, cleaning, treatment or storage of seeds. Islam et al. (2015) reported that OrtT plays a potential role in maintaining bacterial cell fitness during stress. It was reported that TA loci are stress-

response elements in *E. coli* that help cells survive unfavourable growth conditions (Pandey & Gerdes, 2005). Ramisetty & Sonthosh (2017) stated that the range of TAs in bacterial genomes could influence ecological fitness and drive the evolution of bacterial genomes. Toxins may enable adaptation of an organism to changing environments by reducing and preventing cell growth or causing some cells to die (Yamaguchi et al., 2011). The EF hand domain protein (Table 4.3) is the C-terminal activator domain of pesticin (Patzer et al., 2012), a protein that is toxic to bacteria when taken up at the target site in the cell (Cascales et al., 2007; Patzer et al., 2012). Pesticin is a toxin secreted by *Yersinia pestis* and other Gammaproteobacteria, which kills related bacteria occupying the same ecological niche. In addition, it was speculated that pesticin is a bacteriocin that causes hydrolysis of peptidoglycan (Cascales et al., 2007; Patzer et al., 2012). Pesticin production by the onion strain BD 1274 of *P. agglomerans* might contribute to the virulence of this strain on onion. Future studies should investigate the role of the gene encoding pesticin in pathogenicity of *P. agglomerans* to onion.

De Maayer et al. (2014) reported that phages are important for transferring bacterial fitness and pathogenicity factors to their host. A large number of phage sequences detected in strain BD 1212 suggests that phages might play a role in the fitness of non-pathogenic *P. agglomerans*. Studies have showed that such ISs play an important role in a changing the DNA of the cell by allowing the bacterial host cell to adapt to new environmental challenges and to colonise new niches (Mira et al., 2006; Vandecraen et al., 2017). Cesbron et al. (2015) proposed that organisms harbouring ISs are exposed to a variety of mechanism such as horizontal gene transfer that enhance genome plasticity. The differences found in mobile genetic elements identified in the strains of *P. agglomerans* pathogenic and non-pathogenic to onion this study (Tables 2, 3 and S2) confirm a potential role of these elements in the variation or differences of related strains colonising the same habitat (de Maayer et al., 2014).

Other pathogenicity related genes identified in the onion non-pathogenic strain BD 1212 of *P. agglomerans* include genes encoding the probable transcription regulator, antirestriction protein ArdA, homoserine O-succinyltransferase, LOS biosynthesis enzyme LBGB, type 4 and putative colicin immunity protein. The colicin immunity protein is a plasmid immunity protein required for protecting the bacterial cell against colicin (a toxin released by bacteria when they a stressed) (Kleanthous & Walker, 2001). In the genome of the onion pathogenic strain, BD 1274, an uncharacterised protein YfgJ, protein containing domains DUF403, a possible Neuromedin U precursor, a pactoprenol-linked glucose translocase, and a putative cytoplasmic

protein USSDB7A were identified. Furthermore, an osmotically inducible protein C/organic hydroperoxide resistance (OsmC/Ohr) family protein, a putative recombinase, an uncharacterized protein YqeH, the CpmJ protein, a transcription regulator BH0900 (for enhancement of xylanase A production), a putative N-acetylglucosamine kinase, a putative site specific recombinase, a putative glycoporin and a virulence sensor protein bvgS precursor were found in BD 1274. BvgS is a sensor kinase, which enables the bacterium to sense environments and control expression of specific genes (Parkinson & Kofoid, 1992).

4.6 Conclusion

This study compared, for the first time, the genomes of *P. agglomerans* strains that were non-pathogenic and pathogenic on onion that had been isolated from the same onion seed lot. Unique genes were identified that differentiated the non-pathogenic and pathogenic *P. agglomerans* strains. Unique genes present in the pathogenic strain only might play a role in bacterial colonisation, fitness, survival and pathogenicity of that strain. Plant cell wall degrading enzymes and the protein encoding the T3SS in the non-pathogenic strain might be important for colonisation and bacterial fitness in onion plants. In contrast, conjugal transfer might play a major role in the pathogenicity to onion of strain BD 1274 of *P. agglomerans*. In addition, the toxins identified only in the pathogenic strain could play a significant role in the onion virulence of this strain. The sequences of these toxins identified only in the pathogenic strain could be used for designing PCR primers for rapid identification and differentiation of strains of *P. agglomerans* that are pathogenic on onion from those that are not onion pathogens. This could have particularly valuable applications for onion stakeholders, e.g., for designing a molecular onion seed health assay to identify seed lots that contain pathogenic strains of this bacterial species that need to be subjected to an appropriate treatment (e.g., hot water treatment) in order to reduce the risk of seed transmission, to study interactions of pathogenic and non-pathogenic strains of *P. agglomerans* in onion seed or during onion bulb and seed production, or to monitor the efficacy of production practices, environmental conditions, and management recommendations on onion pathogenic strains of *P. agglomerans*. Thus, more genomes of other pathogenic and non-pathogenic strains of *P. agglomerans* must be sequenced to find out if the unique genes identified in the pathogenic BD 1274 strain are present in all strains pathogenic on onion plants.

4.7 References

Alfano, J.R. & Collmer, A. 2004. Type III secretion system effector proteins: double agents in bacterial disease and plant defence. *Annu Rev Phytopathol* 42:385-414.

Alfano, J.R. & Collmer, A. 1997. The type III (Hrp) secretion pathway of plant pathogenic bacteria: trafficking harpins, Avr proteins, and death. *J Bacteriol* 179:5655-5662.

Alikhan, N.F., Petty, N.K., Ben Zakour, N.L. & Beatson, S.A. 2011. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 12:402.

Asis, C.A. & Jr, Adachi, K. 2004. Isolation of endophytic diazotroph *Pantoea agglomerans* and nondiazotroph *Enterobacter asburiae* from sweet potato stem in Japan. *Lett Appl Microbiol 38:19-23.*

Aziz, R.K, Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A. & Zagnitko, O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9:75.

Barash, I. & Manulis-Sasson, S. 2009. Recent evolution of bacterial pathogens: the gall forming *Pantoea agglomerans* case. *Annu Rev Phytopathol* 47:133-152.

Barret, M., Briand, M., Bonneau, S., Préveaux, A., Valière, S., Bouchez, O., Hunault, G., Simoneau, P. & Jacques, M.A. 2015. Emergence shapes the structure of the seed microbiota. *Appl Environ Microbiol* 81:1257-1266.

Barret, M., Guimbaud, J.F., Darrasse, A. & Jacques, M.A. 2016. Plant microbiota affects seed transmission of phytopathogenic microorganisms. *Mol Plant Pathol* 17:791-795.

Bender, C.L., Alarcon-Chaidez, F. & Gross, D.C. 1999. *Pseudomonas syringae* phytotoxins: mode of action, regulation and biosynthesis by peptide and polyketide synthetases. *Microbiol Mol Biol Rev* 63:266-292.

Berne, C., Ducret, A., Hardy, G.G. & Brun. V.Y. 2015. Adhesins involved in attachment to abiotic surfaces by Gram negative bacteria. *Microbiol Spectr* 3(4).

Beranek, A. Zettl, M., Lorenzoni, K., Schauer, A., Manhart, M. & Koraimann, G. 2004. Thirty-eight C-terminal amino acids of the coupling protein TraD of the F-Like conjugative resistance plasmid R1 are required and Sufficient to confer binding to the substrate selector protein TraM. *J Bacteriol* 186:6999-7006.

Blom, J., Albaum, S.P., Doppmeier, D., Pühler, A., Vorhölter, F-J., Zakrzewski, M. &

Goesmann, A. 2009. EDGAR: a software framework for the comparative analysis of prokaryotic genomes. *BMC Bioinformatics* 10:154.

Blot, N., Berrier, C., Hugouvieux-Cotte-Pattat, N., Ghazi, A. & Condemine, G. 2002. The oligogalacturonate-specific porin KdgM of *Erwinia chrysanthemi* belongs to a new porin family. *J Biol Chem* 277:7936-7944.

Bolger, A.M., Lohse, M. & Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30:2114-2120.

Bonas, U. & Lahaye, T. 2002. Plant disease resistance triggered by pathogen derived molecules: refined models of specific recognition. *Curr Opin Microbiol* 5:44-50.

Brader, G., Compant, S., Vescio, K., Mitter, B., Trognitz, F., Ma, L-J. & Sessitsch, A. 2017. Ecology and genomic insights into plant-pathogenic and plant-nonpathogenic endophytes. *Annu Rev Phytopathol* 55:61-83.

Brettin, T., Davis, J.J., Disz, T., Edwards, R.A., Gerdes. S., Olsen, G.J., Olson, R., Overbeek, R., Parrello, B., Pusch, G.D., Shukla, M., Thomason, J.A 3rd., Stevens, R., Vonstein, V., Wattam, A.R. & Xia, F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Scientific reports 5: 8365.

Buttimer, C., McAuliffe, O., Ross, R.P., Hill, C., O'Mahony, J. & Coffey, A. 2017. Bacteriophages and bacterial plant diseases. *Front Microbiol* 8:34.

Büttner, D. 2012. Protein export according to schedule: architecture, assembly, and regulation of type III secretion systems from plant and animal pathogenic bacteria. *Microbiol Mol Biol Rev* 76:262-310.

Cascales, E., Buchanan, S.K., Duché, D., Kleanthous, C., Lloubès, R., Postle, K., Riley, M., Slatin, S. & Cavard, D. 2007. Colicin biology. *Microbiol Mol Biol Rev* 71:158-229.

Cesbron, S., Briand, M., Essakhi, S., Gironde, S., Boureau, T., Manceau, C., Fischer-Le Saux, M. & Jacques, M-A. 2015. Comparative genomics of pathogenic and nonpathogenic strains of *Xanthomonas arboricola* unveil molecular and evolutionary events linked to pathoadaptation. *Front Plant Sci* 6:1126.

Christendat, D., Saridakis, V., Dharamsi, A., Bochkarev, A., Pai, F.E., Arrowsmith, H.C. & Edwards, M.A. 2000. Crystal Structure of dTDP-4-keto-6-deoxy-D-hexulose 3,5-Epimerase from *Methanobacterium thermoautotrophicum* Complexed with dTDP. *J Biol Chem* 275:24608-24612.

Christie, P.J. 2016. The mosaic type IV secretion systems. *EcoSalPlus* 7. doi:10.1128.

Christie, E.G. & Calendar, R. 2016. Bacteriophage P2. *Bacteriophage* 6(1):e114578

Christie, P.J. & Cascales, E. 2005. Structural and dynamic properties of bacterial type IV secretion systems (review). *Mol Membr Biol* 22:51-61.

Christie, P.J., Atmakuri, K., Krishnamoorthy, V., Jakubowski, S. & Cascales, E. 2005. Biogenesis, architecture, and function of bacterial type IV secretion systems. *Annu Rev Microbiol* 59:451-485.

Cruz, A.T., Cazacu, A.C. & Allen, C.H. 2007. *Pantoea agglomerans*-a plant pathogen causing human disease. *J Clin Microbiol* 45:1989-1992.

Darling, A.C., Mau, B., Blattner, F.R. & Perna, N.T. 2004. Mauve: multiple alignment of conserved genomic sequences with rearrangements. *Genome Res* 14:1394-1403.

Darsonval, A., Darrasse, A., Meyer, D., Demarty, M., Durand, K., Bureau, C., Manceau, C. & Jacques, M.A. 2008. The Type III secretion system of *Xanthomonas fuscans* subsp. *fuscans* is involved in the phyllosphere colonisation process and in transmission to seeds of susceptible beans. *Appl Environ Microbiol* 74:2669-2678.

Darsonval, A., Darrasse, A., Durand, K., Bureau, C., Cesbron, S. & Jacques, M.A. 2009. Adhesion and fitness in the bean phyllosphere and transmission to seed of *Xanthomonas fuscans* subsp. *fuscans*. *Mol Plant-Microbe Interact* 22:747-757.

de Maayer, P., Chan, W.Y., Rubagotti, E., Venter, S.N., Toth, I.K., Birch, P.R.J. & Coutinho, T.A. 2014. Analysis of the *Pantoea ananatis* pan-genome reveals factors underlying its ability to colonize and interact with plant, insect and vertebrate hosts. *BMC Genomics* 15:404.

Diepold, A. & Armitage, P.J. 2015. Type III secretion systems: the bacterial flagellum and the injectisome. *Phil Trans R Soc* B 370.

Doughari, J.H. 2015. An overview of plant immunity *J Plant Pathol Microbiol* 6:322.

Edens, D.G., Gitaitis, R.D., Sanders, F.H. & Nischwitz, C. 2006. First report of *Pantoea agglomerans* causing a leaf blight and bulb rot of onions in Georgia. *Plant Dis* 90:1551.

Evans, T.J., Coulthurst, S.J., Komitopoulou, E. & Salmond, G.P.C. 2010. Two mobile *Pectobacterium atrosepticum* prophages modulate virulence. *FEMS Microbiol Lett* 304:195-202.

Feng, Y., Shen, D. & Song, W. 2006. Rice endophyte *Pantoea agglomerans* YS19 promotes host plant growth and affects allocation of host photosynthesis. *J Appl Microbiol* 100:938-945.

González, D.A., Franco, M.A., Contreras, N., Galindo-Castro, I. & Jayaro, J. 2015. First report of *Pantoea agglomerans* causing rice leaf blight in Venezuela. *Plant Dis* 99:552.

Goszczynska, T., Moloto, V.M., Venter, S.N. & Coutinho, T.A. 2006. Isolation and

identification of *Pantoea ananatis* from onion seed in South Africa. *Seed Science Technol* 34:677-690.

Grainge, I & Sherratt, J.D. 1999. Xer Site-specific Recombination: DNA strand rejoining by recombinase. *J Biol Chem* 274:6763-6769.

Grijpstra, J., Arenas, J., Rutten, L. & Tommassen, J. 2013. Autotransporter secretion: varying on a theme. *Res Microbiol* 164:562-582.

Grohmann, E., Peter, J., Christie, J.P., Waksman, G. & Backert, S. 2018. Type IV secretion in Gram negative and Gram positive bacteria. *Mol Microbiol* 107:455-471.

Haiko, J. & Westerlund-Wikström, B. 2013. The role of the bacterial flagellum in adhesion and virulence. *Biol* 2:1242-1267.

Hardoim, P.R., Hardoim, C.C.P., van Overbeek, L.S. & van Elsas, J.D. 2012. Dynamics of seedborne rice endophytes on early plant growth stages. *PLoS ONE* 7:e30438.

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

Henderson, I.R., Navarro-Garcia, F., Desvaux, M. Fernandez, R.C. & Ala'Aldeen, D. 2004. Type V protein secretion pathway: the autotransporter story. *Microbiol Mol Biol Rev* 68:692-744.

Hutter, A.J.C., Lehner, R., Wirth, C., Condemine, G., Peneff, C. & Schirmer, T. 2014. structure of the oligogalacturonate-specific KdgM porin. *Acta Cryst* 70:1770-1778.

Imada, K., Minaminob, T., Uchidaa, Y., Kinoshitaa, M. & Namba, K. 2016. Insight into the flagella type III export revealed by the complex structure of the type III ATPase and its regulator. *PNAS* 113:3633-3638.

Islam, S., Benedik J.M. & Wood K.T. 2015. Orphan Toxin OrtT (YdcX) of *Escherichia coli* reduces growth during the stringent response. *Toxin* 7:299-321.

Johnston-Monje, D. & Raizada, M.N. 2011. Conservation and diversity of seed associated endophytes in *Zea* across boundaries of evolution, ethnography and ecology. *PLoS ONE* 6(6): e20396.

Johnson, K.L. & Walcott, R.R. 2013. Quorum sensing contributes to seed-to seedling transmission of *Acidovorax citrulli* on watermelon. *J Phytopathol* 161:562-573.

Jørgensen, M.G., Pandey, D.P., Jaskolska, M. & Gerdes, K. 2009. HicA of *Escherichia coli* defines a novel family of translation-independent mRNA interferases in bacteria and archaea. *J Bacteriol* 1191-1199.

Kleanthous, C. & Walker, D. 2001. Immunity proteins: enzyme inhibitors that avoid the active site. *Trends Biochem Sci* 26:624-631.

Korotkov, K.V., Sandkvist, M. & Hol, W.G.J. 2012. The type II secretion system: biogenesis, molecular architecture and mechanism. *Nat Rev Microbiol* 10:336-351.

Lawley, T.D., Klimke, W.A., Gubbins, M.J. & Frost, L.S. 2003. F factor conjugation is a true type IV secretion system. *FEMS Microbiol Lett* 224:1-15.

Lindh, E., Kjaeldgaard, P., Frederiksen, W. & Ursing, J. 1991. Phenotypical properties of *Enterobacter agglomerans* (*Pantoea agglomerans*) from human, animal and plant sources. *APMIS* 99:347-352.

Liu, W-Y., Wong, C-F., Chung, K.M-K., Jiang, J-W. & Leung F.C-C. 2013. Comparative genome analysis of *Enterobacter cloacae*. *PLoS ONE* 8(9): e74487.

Lòpez-Fernàndez, S., Sonego, P., Moretto, M., Pancher, M., Engelen, K., Pertot, I. & Campisano, A. 2015. Whole-genome comparative analysis of virulence genes unveils similarities and differences between endophytes and other symbiotic bacteria. *Front Microbiol* 6:419.

Makarova, K.S., Grishin, N.V. & Koonin, E.V. 2006. The HicAB cassette, a putative novel, RNA-targeting toxin-antitoxin system in archaea and bacteria. *Bioinformatics* 22:2581-2584.

Manulis, S. & Barash, I. 2003. *Pantoea agglomerans* pvs. *gypsophilae* and *betae*, recently evolved pathogens? *Mol Plant Pathol* 4:307-314.

Marguerettaz, M., Pieretti, I., Gayral, P., Puig, J., Brin, C., Cociancich, S., Poussier, S., Rott, P. & Royer, M. 2011. Genomic and evolutionary features of the SPI-1 type III secretion system that is present in *Xanthomonas albilineans* but is not essential for xylem colonization and symptom development of sugarcane leaf scald. *Mol. Plant Microbe Interact.* 24, 246-259.

Medrano, E.G. & Bell, A.A. 2007. Role of *Pantoea agglomerans* in opportunistic bacterial seed and boll rot of cotton (*Gossypium hirsutum*) grown in the field. *J Appl Microbiol* 102:134-143.

Melotto, M. & Kunkel, B.N. 2013. Virulence strategies of plant pathogenic bacteria. In: The Prokaryotes - Prokaryotic Physiology and Biochemistry (Rosenberg, E., ed.), pp. 6175. (Texas, USA; Springer).

Mira, A., Pushker, R. & Rodríguez-Valera, F. 2006 .The Neolithic revolution of bacterial genomes. *Trends Microbiol* 14:200-206.

Moloto, V.M., Goszczynska, T., du Toit, L.J. & Coutinho, T.A. 2017. A new pathovar of *Pseudomonas syringae*, pathovar *allii*, isolated from onion plants exhibiting symptoms of blight. *Eur J Plant Pathol* 147:591-603.

Morales-Valenzuela, G., Silva-Rojas, V.H., Ochoa-Martínez, D., Valadez-Moctezuma, E., Alarcón-Zúñiga, B., Zelaya-Molina, X.L., Córdova-Téllez, L., Mendoza-Onofre, L., H. Vaquera-Huerta, H., Carballo-Carballo, A., Farfán-Gómez, A. & Ávila-Quezada, G. 2007. First Report of *Pantoea agglomerans* causing leaf blight and vascular wilt in maize and sorghum in Mexico. *Plant Dis* 91:1365.

Moretti, C., Cortese, C., Passos da Silva, D., Venturi, V., Torelli, E., Firrao, G. & Buonaurio, R. 2014a. Draft genome sequence of a hypersensitive reaction inducing *Pantoea agglomerans* strain isolated from olive knots caused by *Pseudomonas savastanoi* pv. *savastanoi*. *Genome Announc* 2:e00774-14.

Moretti, C., Cortese, C., Passos da Silva, D., Venturi, V., Firrao, G. & Buonaurio, R. 2014b. Draft genome sequence of *Erwinia oleae*, a bacterium associated with olive knots caused by *Pseudomonas savastanoi* pv. *savastanoi*. *Genome Announc* 2:e01308-14.

Mundt, J.O. & Hinkle, N.F. 1976. Bacteria within ovules and seeds. *Appl Environ Microbiol* 32:694-698.

Nilsson, E.S & Haggård-Ljungquist, E. 2007. Evolution of P2-like phages and their impact on bacterial evolution. *Res Microbiol* 158:311-317.

Nunes, C., Usall, J., Teixidó, N., Fons, E. & Vinãs, I. 2002. Post-harvest biological control by *Pantoea agglomerans* (CPA-2) on Golden Delicious apples. *J Appl Microbiol* 92:247-255.

Palmer, M., Steenkamp, E.T., Coetzee, M.P.A., Blom, J. & Venter, S.N. 2018. Genome-based characterisation of biological processes that differentiate closely related bacteria. *Front Microbiol* 9:113.

Pandey, D.P. & Gerdes, K. 2005. Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Res* 33:966-976.

Parkinson, J.S. & Kofoid, E.C. 1992. Communication modules in bacterial signalling proteins. *Annu Rev Genet* 26:71-112.

Patzer, I.S., Albrecht, R., Braun, V. & Zeth, K. 2012. Structural and mechanistic studies of pesticin, a bacterial homolog of phage lysozymes. J Biol Chem 287:23381-23396.

Pfeilmeier, S., Caly, D.L. & Malone, J.G. 2016. Bacterial pathogenesis of plants: future challenges from a microbial perspective. *Mol Plant Pathol* 17:1298-1313.

Pizarro-Cerdá, J. & Cossart, P. 2006. Bacterial adhesion and entry into host cells. *Cell* 124: 715.

Plumbridge, J. & E. Vimr, E. 1999. Convergent pathways for utilization of the amino sugars

N-acetylglucosamine, N-acetylmannosamine, and N-acetylneuraminic acid by *Escherichia coli*. *J Bacteriol* 181:47-54.

Preston, G.M., Bertrand, N. & Rainey, P.B. 2001. Type III secretion in plant growth promoting *Pseudomonas fluorescens* SBW25. *Mol Microbiol* 41:999-1014.

Rahim, R., Ochsner, U.A., Olvera, C., Graninger, M., Messner P., Lam, J.S. & Soberón-Chávez, G. 2001. Cloning and functional characterisation of the *Pseudomonas aeruginosa* rhlC gene that encodes rhamnosyltransferase 2, an enzyme responsible for di-rhamnolipid biosynthesis. *Mol Microbiol* 40:708-718.

Ramisetty, C.M.B. & Santhosh, S.R. 2017. Endoribonuclease type II toxin–antitoxin systems: functional or selfish? *Microbiol* 163:931-939.

Renzi, F., Caffarelli, E., Laneve, P., Bozzoni, I., Brunori, M. & Vallone, B. 2006. The structure of the endoribonuclease XendoU: From small nucleolar RNA processing to severe acute respiratory syndrome coronavirus replication. *PNAS USA* 103:12365-12370.

Rezzonico, F., Binder, C., Défago, G. & Moënne-Loccoz, Y. 2005. The type III secretion system of biocontrol *Pseudomonas fluorescens* KD targets the phytopathogenic Chromista *Pythium ultimum* and promotes cucumber protection. *Mol Plant-Microbe Interact* 18:991-1001.

Sibanda S., Kwenda, S., Tanui, K.C. Shyntum, Y.D., Coutinho, A.T. & Moleleki, N.L. 2018. Transcriptome profiling reveals the EanI/R Quorum sensing regulon in *Pantoea ananatis* LMG 2665T. *Genes* 9:148.

Shariati, V., Malboobi, M.A., Tabrizi, Z., Tavakol, E., Owilia, P. & Safari, M. 2017. Comprehensive genomic analysis of a plant growth-promoting rhizobacterium *Pantoea agglomerans* strain P5. *Scientific Reports* 7:15610.

Siguier, P., Gourbeyre, E. & Chandler, M. 2014. Bacterial insertion sequences: their genomic impact and diversity. *FEMS Microbiol Rev* 38:865-891.

Souza, D.P., Oka, G.U., Alvarez-Martinez, C.E., Bisson-Filho, A.W., Dunger, G., Hobeika, L., Cavalcante, N.S., Alegria, M.C., Barbosa, L.R., Salinas, R.K., Guzzo, C.R. & Farah, C.S. 2015. Bacterial killing via a type IV secretion system. *Nat Commun* 6:6453.

Tampakaki, A.P. 2014. Commonalities and differences of T3SSs in rhizobia and plant pathogenic bacteria. *Front Plant Sci* 5:114.

Toth, I.K., Bell, K.S., Holeva, M.C. & Birch, P.R. 2003. Soft rot erwiniae: from genes to genomes. *Mol Plant Pathol* 4:17-30.

Truyens, S., Weyens, N., Cuypers, A. & Vangronsveld, J. 2015. Bacterial seed endophytes:

genera, vertical transmission and interaction with plants. *Environ Microbiol Rep* 7:40-50.

Vandecraen, J., Chandler, M., Aertsen, A. & Van Houdt, R. 2017. The impact of insertion sequences on bacterial genome plasticity and adaptability. *Crit Rev Microbiol* 43:709-730.

Van Melderen, L. & de Bast, M.S. 2009. Bacterial toxin-antitoxin systems: More than selfish entities? *PLoS Genet* 5(3): e1000437.

Vo, J.L., Martnez Ortiz, C.G., Subedi, P., Keerthikumar, S., Mathivanan, S., Paxman J.J. & Heras, B. 2017. Autotransporter adhesins in *Escherichia coli* pathogenesis. *Proteomics* 17:1600431.

Vonstein, V., Warren, A., Will, R., Wilson, M.J., Yoo, H.S., Zhang, C., Zhang, Y. & Sobral, B.W. 2014. PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Res* 42:D581-591.

Vorhölter, F.J., Wiggerich, H.G., Scheidle, H., Sidhu, V.K., Mrozek, K., Küster, H., Pühler, A. & Niehaus, K. 2012. Involvement of bacterial TonB-dependent signalling in the generation of an oligogalacturonide damage-associated molecular pattern from plant cell walls exposed to *Xanthomonas campestris* pv. *campestris* pectate lyases. *BMC Microbiol* 12:239.

Wattam, A.R., Abraham, D., Dalay, O., Disz, T.L., Driscoll, T., Gabbard, J.L., Gillespie, J.J., Gough, R., Hix, D., Kenyon, R., Machi, D., Mao, C., Nordberg, E.K., Olson, R., Overbeek, R., Pusch, G.D., Shukla, M., Schulman, J., Stevens, R.L., Sullivan, D.E., Sullivan, D.E., Vonstein, V., Warren, A., Will, R., Wilson, M.J., Yoo, H.S., Zhang, C., Zhang, Y. & Sobral, B.W. 2014. PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Res* 42:D581-591.

Weller-Stuart, T., Toth, I., de Maayer. P. & Coutinho, T. 2017. Swimming and twitching motility are essential for attachment and virulence of *Pantoea ananatis* in onion seedlings. *Mol Plant Pathol* 18:734-745.

Yamaguchi, Y., Park J-H. & Inouye, M. 2011. Toxin-antitoxin systems in bacteria and archaea. *Annu Rev Genet* 45:61-79.

Załuga, J., Stragier, P., Baeyen, S., Haegeman, A., Van Vaerenbergh, J., Maes, M. & De Vos, P. 2014. Comparative genome analysis of pathogenic and non-pathogenic *Clavibacter* strains reveals adaptations to their lifestyle. *BMC Genomics* 15:392.

Zechner, E.L., Lang, S. & Schildbach, J.F. 2012. Assembly and mechanisms of bacterial type IV secretion machines. *Philos Trans R Soc B* 367:1073-1087.

Zhang, D., Iyer, M.L. & Aravind, L. 2011. A novel immunity system for bacterial nucleic acid degrading toxins and its recruitment in various eukaryotic and DNA viral systems. *Nucleic Acids Res* 39:4532-4552.

4.8 Tables

Table 4.1 Comparison of genome characteristics based on PATRIC annotation (Wattam et al., 2014) of strains of *Pantoea agglomerans* isolated from onion seed that were non-pathogenic (BD 1212) and pathogenic (BD 1274) on onion, respectively.

Characteristics	***P. agglomerans* BD 1212**	***P. agglomerans* BD 1274**
Genome size (bp)	4, 875 404	4, 968 508
DNA G+C (bp)	55.1	55
Total number of genes	4751	4918
Protein coding genes	4492	4568
rRNA	4	11
tRNA	52	53
Pseudogenes	0	0
Genes with hypothetical proteins	942	1036
Genes with function prediction	3809	3882
Genes with pathway assignment	879	905

Table 4.2 Genes present only in strain BD 1212 of *Pantoea agglomerans* that is non-pathogenic on onion plants, and not in strain BD1274 that is pathogenic on onion.

Categories	Name of gene	Contig	Size (bp)
Carbohydrate			
Mannose metabolism	Mannose-1-phosphate guanylyltransferase	43	471
Insertion element (transposase)	IS1 related to InsA	4	90
	ISSod13, transposase	7	346
	IS66 transposase family	7	478
	IS66 transposase family	7	83
	Transposase InsF for insertion sequence IS3 @ Mobile element protein	7	186
Cell wall degrading enzyme	Oligogalacturonate-specific porin	3	246
T3SS	Type III secretion protein HrpJ	9	380
	Type III secretion inner membrane channel protein (LcrD,HrcV,EscV,SsaV, homologous to flagellar components)	9	707
	Type III secretion cytoplasmic ATP synthase (YscN,SpaL,MxiB,HrcN,EscN)	9	452
	Type III secretion spans bacterial envelope protein (YscO)	9	154
	Type III secretion protein HrcQb	9	347
	Type III secretion inner membrane protein (YscR,SpaR,HrcR,EscR, homologous to flagellar export components)	9	217
	Type III secretion inner membrane protein (YscS, homologous to flagellar export components)	9	86
	Type III secretion inner membrane protein (YscT,HrcT,SpaR,EscT,EpaR1, homologous to flagellar export components)	9	251
	Type III secretion inner membrane protein (YscU,SpaS,EscU,HrcU,SsaU, homologous to flagellar export components)	9	361
	Type III secretion outermembrane pore forming protein (YscC,MxiD,HrcC,InvG)	9	682
	Type III secretion protein	9	198
	HrpD	9	193

	Type III secretion bridge between inner and outermembrane lipoprotein (YscJ,HrcJ,EscJ,PscJ)	9	254
	HrpB	9	122
Non-fimbrial and fimbrial adhesins	Putative fimbrial structural subunit	7	148
	Adhesin autotransporter YdeU	7	64
	Putative fimbrial structural subunit	7	14
	AidA-I adhesin-like protein	7	543
	Adhesin autotransporter YdeU	94	174
	Uncharacterized protein YcgV (AiDA)	68	951
	Filamentous haemagglutinin family outer membrane protein	100	522
Bacteriophage/Phage	FIG118045: Phage immunity repressor protein	1	181
	FIG070121: Phage capsid and scaffold protein	1	238
	Bacteriophage CI repressor	19	287
	Phage protein	19	125
	Regulatory protein CII	19	169
	Phage replication protein	19	751
	UV induction of prophage	19	61
	Gene D protein	61	393
	Phage-related tail protein	61	175
	Phage protein	61	820
	Putative phage tail protein	61	92
	Phage major tail tube protein	61	169
	Phage tail sheath monomer	61	389
	Phage tail fiber protein	65	205
	Phage tail fibers	65	202
	Baseplate assembly protein J	65	302
	Phage baseplate assembly protein	65	116
	Baseplate assembly protein V	65	196
	Phage tail completion protein	65	149
	Phage tail protein	65	155
	Prophage lysozyme; Phage lysin	65	169
	Phage tail completion protein	65	67
	Phage head completion-stabilization protein	65	156
	Phage terminase, endonuclease subunit	65	222

	Phage major capsid protein	65	370
	Phage capsid scaffolding protein	65	282
	Phage terminase, ATPase subunit	65	587
	Phage-related capsid packaging protein	65	344
	Phage T7 exclusion protein associated hypothetical protein	67	250
	Phage T7 exclusion protein	67	594
Other genes	Tellurite resistance protein-related protein	4	222
	Uncharacterized protein YaiN in in formaldehyde detoxification operon	8	91
	COG2183: Transcriptional accessory protein	10	97
	Ferric iron ABC transporter, iron-binding protein	11	341
	COG2183: Transcriptional accessory protein	18	99
	Succinate-acetate/proton symporter SatP	2	189
	Probable transcription regulator protein	4	65
	Probable integral membrane protein NMA1777	4	297
	Putative MutT/nudix-family hydrolase	5	157
	Prolyl endopeptidase	5	128
	Antirestriction protein ArdA	7	166
	UPF0401 protein YkfF	7	77
	Plasmid SOS inhibition protein PsiB	7	145
	Homoserine O-succinyltransferase	24	309
	LOS biosynthesis enzyme LBGB	52	153
	TonB-dependent hemin, ferrichrome receptor	66	790
	Hemin transport protein HmuS	66	343
	Putative deoxyribonuclease similar to YcfH, type 4	67	237
	RNA:NAD 2'-phosphotransferase	77	180
	Putative colicin immunity protein	78	89
	Phytochrome-like protein; Cph2	99	321

Table 4.3. Genes present only in strain BD 1274 of *Pantoea agglomerans* that is pathogenic on onion plants, and not in strain BD1212 that is non-pathogenic on onion.

Categories	Name of gene	Contig	Size
Carbohydrate			
Pentose phosphate pathway	Ribose 5-phosphate isomerase B	18	152
Sucrose metabolism	Sucrose permease, major facilitator superfamily	78	413
Glycolysis and Gluconeogenesis	Xylan 1,4-beta-xylosidase	78	503
	unsaturated glucuronyl hydrolase	78	392
Glyoxylate and dicarboxylate metabolism	Oxalate decarboxylase	234	408
	Putative N-acetylglucosamine kinase	133	44
Nucleotide	dTDP-4-dehydrorhamnose 3,5-epimerase	79	182
	dTDP-4-dehydrorhamnose reductase	79	295
	deoxycytidylate deaminase-related protein	224	542
Insertion element sequence/transposase	Putative transposase InsK for insertion sequence element IS150	7	189
	Transposase ISL3	135	132
	Transposase InsO for insertion sequence element IS911	154	57
T4SS	TraE family protein	5	202
	TraB	5	462
	TraW	5	213
	TraU	5	330
	TrbC	5	209
	TraN	5	616
	TraF	5	265
	TrbB	5	207
	TraH	5	455
	TraG	5	1031
	TraD	5	867

	TraI	5	1951
Toxin	Phage DNA binding protein CopG (ACLAME 166)	2	141
	Probable mRNA interferase HicA	2	60
	EF hand domain protein	116	221
	YefM protein (antitoxin to YoeB)	150	251
Bacteriophage/phage	Putative bacteriophage protein	110	39
	Lambdoid phage Rac integrase	110	401
Vitamin	Biotin synthesis protein BioC	164	770
Other genes	Uncharacterized protein YfgJ	1	73
	Protein containing domains DUF404, DUF407	4	478
	Protein containing domains DUF403	4	308
	Protein containing transglutaminase-like domain, putative cysteine protease	4	260
	Thiol:disulfide interchange protein DsbG precursor	4	253
	Gene II and X proteins	4	137
	Putative transmembrane protein	4	142
	Sulfide:quinone oxidoreductase, Type II	4	410
	ParA-like protein	4	341
	Copper resistance protein B	4	256
	Putative alkyl/aryl-sulfatase YjcS	7	658
	Possible Neuromedin U precursor	7	263
	Bactoprenol-linked glucose translocase	14	118
	Arginase	30	305
	Putative cytoplasmic protein USSDB7A	32	96
	ADP-ribosylglycohydrolase	56	363
	OsmC/Ohr family protein	61	159
	putative recombinase	61	125
	Uncharacterized protein YqeH	67	152
	GTPase	128	108

CpmJ protein	132	178
Transcription regulator (enhancement of xylanase A production) BH0900	132	277
putative site specific recombinase	154	405
Putative glycoporin	175	439
Virulence sensor protein bvgS precursor	189	292
MmgE/PrpD family protein	185	249
Molybdopterin biosynthesis MoeB protein	204	373
UDP-sulfoquinovose synthase	222	388

4.9 Figures

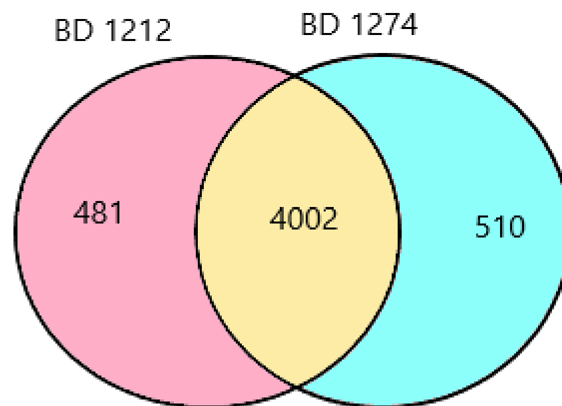


Figure 4.1 A Venn diagram showing shared and unique genes between *Pantoea agglomerans* strains BD 1212 (non-pathogenic on onion) and BD 1274 (pathogenic on onion) isolated from onion seed. The two genomes share 4002 gene families but differ based on 481 unique genes to BD 1212 and 510 genes unique to BD1274. The EDGAR (Efficient Database framework for comparative Genome analyses using BLAST score Ratios) (Blom et al., 2009) web server was used to construct the Venn diagram.

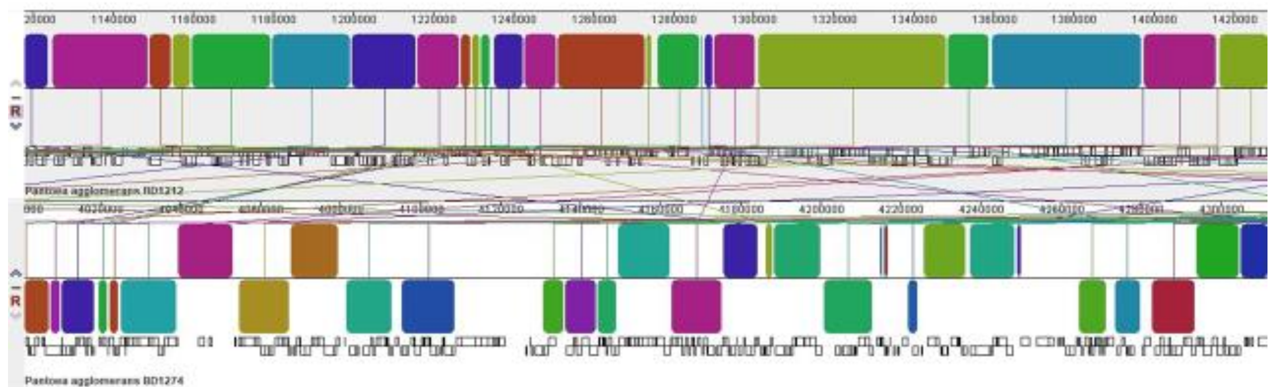


Figure 4.2 Genome alignment showing syntenic blocks between *P. agglomerans* strains BD 1212 (non-pathogenic on onion) (top) and BD 1274 (pathogenic on onion) genomes (bottom) aligned using Mauve v2.3.1 (Darling et al., 2004). Homologous blocks representing annotated proteins among the strains are marked by the same coloured blocks, while gaps correspond to non-homologous regions. Small blocks below the coloured block represent the genes present in different contigs.

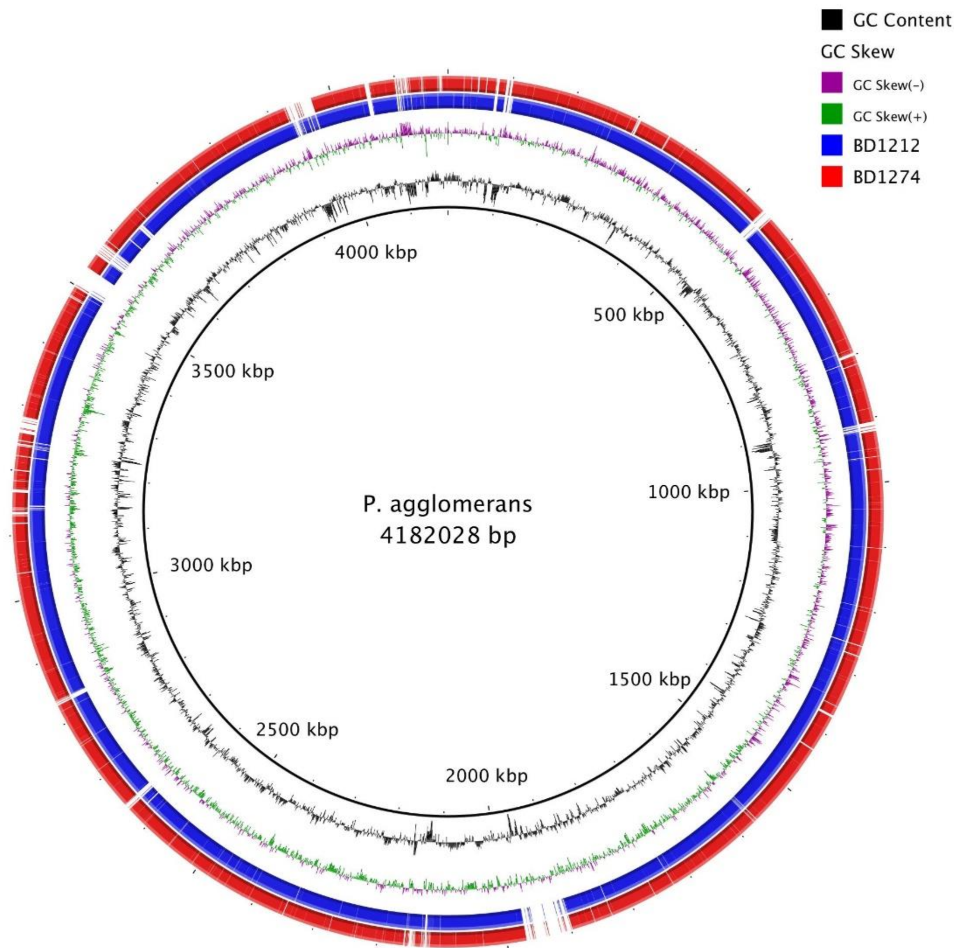


Figure 4.3. Comparison of the circular genome structure of *Pantoea agglomerans* strain BD 1212 (non-pathogenic on onion) and BD 1274 (pathogenic on onion) with a complete reference genome of *P. agglomerans* C410P1 chromosome (NZ_CP016889.1) using the BLAST Ring Image Generator (BRIG v0.95, Alikhan et al., 2011). The inner circle shows the scale (bp). The first and second rings show the GC content (black) and GC skew (green/purple), respectively, with respect to the reference genome. The third and fourth rings show BLAST comparison of BD 1212 (blue) and BD 1274 (red).

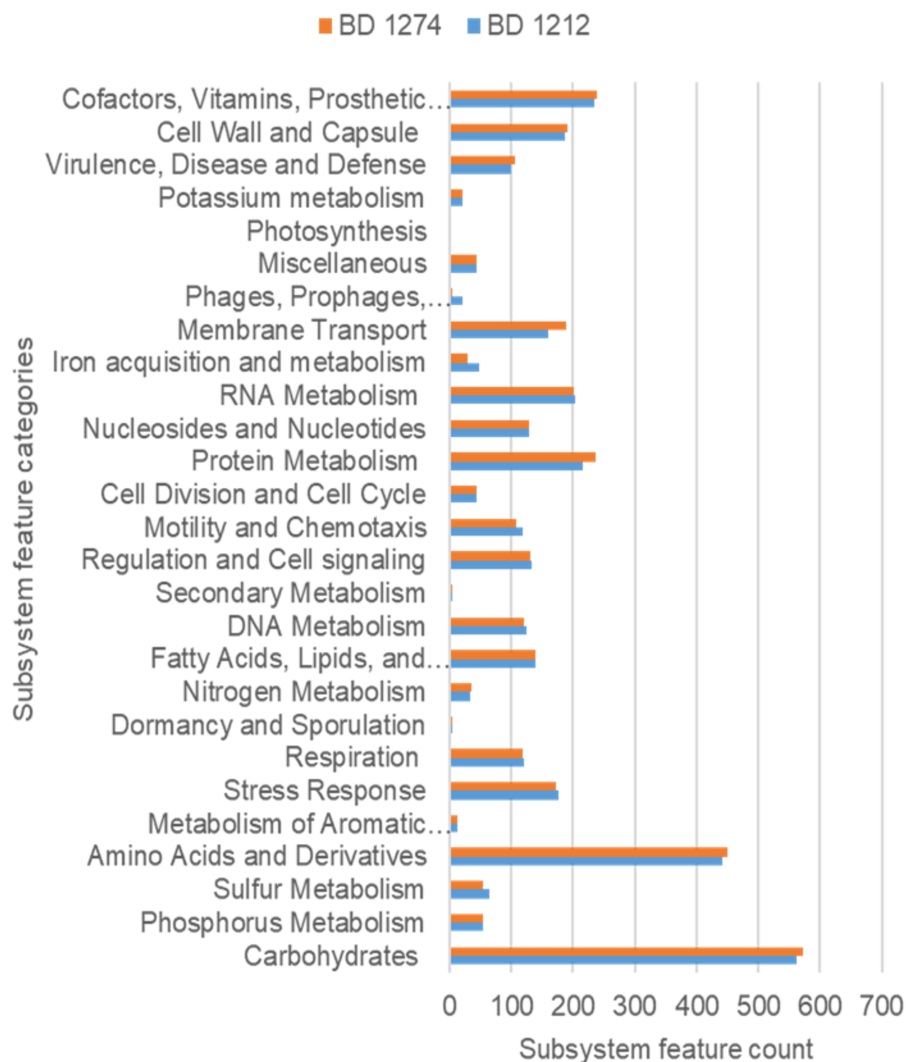


Figure 4.4 Comparison of functional categories of genes between strain BD 1212 (non-pathogenic on onion) and BD 1274 (pathogenic onion) of *Pantoea agglomerans* isolated from onion seed, based on RAST annotation (Aziz et al., 2008). Genes associated with carbohydrate metabolism were most prevalent, followed by genes involved in production of amino acids and derivatives, protein metabolism, cofactors, vitamin and prosthetic, RNA metabolism, cell wall and capsule, stress response, fatty acids, lipids and isoprenoids, regulation and cell signalling and nucleosides and nucleotides.

4.10 Supplementary Information

Table S1. Number and proportion of genes associated with 25 Clusters of Orthologous Groups (COGs) functional categories detected in the genomes of strains BD 1212 (non-pathogenic on onion) and BD 1274 (pathogenic on onion) of *Pantoea agglomerans* isolated from onion seeds.

BD 1212 (non-pathogenic)			**BD 1274 (pathogenic)**		
Code	**Value**	**% of total[a]**	**Value**	**% total[a]**	**Description**
D	44	0.97	47	1.03	Cell cycle control, cell division, chromosome partitioning
M	260	5.79	252	5.52	cell wall/membrane/envelope biogenesis
N	114	2.54	108	2.36	Cell motility
O	139	3.09	153	3.35	Post-translational modification, protein turnover, chaperones
T	243	5.41	240	5.25	Signalling transduction mechanisms
U	113	2.51	96	2.10	Intracellular trafficking, secretion and vesicular transport
V	56	1.24	54	1.18	Defence mechanisms
W	0	0	0	0	Extracellular structure
Y	0	0	0	0	Nuclear structure
Z	0	0	0	0	Cytoskeleton
A	1	0.02	1	0.02	RNA processing and modification
B	0	0	0	0	Chromatin structure and dynamics
J	199	4.43	197	4.31	Translation, ribosomal structure and biogenesis
K	380	8.45	395	8.65	Transcription
L	132	2.93	159	3.48	Replication, recombination and repair
C	207	4.61	222	4.86	Energy production and conversion
E	414	9.21	427	9.35	Amino acid transport and metabolism
F	98	2.18	102	2.23	Nucleotide transport and metabolism
G	416	9.26	448	9.81	Carbohydrate transport and metabolism
H	165	3.67	168	3.68	Coenzyme transport and metabolism
I	118	2.72	121	2.65	Lipid transport and metabolism
P	251	5.59	250	5.47	Inorganic ion transport and metabolism
Q	78	1.74	83	1.82	Secondary metabolites biosynthesis, transport and catabolism
R	497	11.0	504	11.03	General function prediction only
S	389	8.66	408	8.93	Function unknown
[b]-	178	3.96	133	2.91	Not in COGs

[a] The total is based on the total number of the predicted protein coding genes in the annotated genomes.

[b] The number of genes that are not present in COGs.

Table S2. Genes shared between the BD 1212 (non-pathogenic on onion) and BD 1274 (pathogenic on onion) strains of *Pantoea agglomerans* isolated from onion seed.

Categories	Protein encoding genes
Enzymes associated with plant cell wall degradation	Cytoplasmic alpha-amylase
	Glucoamylase
	Glucoamylase
	beta-1,4-glucanase (cellulase)
	Endo-1,4-beta-xylanase
	Endo-1,4-beta-xylanase
	Putative pectin degradation protein
	Polygalacturonase
	Polygalacturonase
	Metalloprotease
	Extracellular metalloprotease precursor
	Exported zinc metalloprotease YfgC precursor
	dTDP-glucose 4,6-dehydratase
	dTDP-glucose 4,6-dehydratase
	Glucose-1-phosphate thymidylyltransferase
	dTDP-fucosamine acetyltransferase
	dTDP-4-amino-4,6-dideoxygalactose transaminase
	Phosphoglucosamine mutase
	UDP-N-acetylglucosamine 2-epimerase
	UDP-N-acetyl-D-mannosamine dehydrogenase
	Alpha-L-Rha alpha-1,3-L-rhamnosyltransferase
	Inner membrane protein YhjD
	LysR family transcriptional regulator YhjC
	Protein YhjJ, putative peptidase
	Uncharacterized protein YhjG
	2-Keto-3-deoxy-D-manno-octulosonate-8-phosphate synthase
	2-Keto-3-deoxy-D-manno-octulosonate-8-phosphate synthase
	3-deoxy-D-manno-octulosonate 8-phosphate phosphatase
	3-deoxy-manno-octulosonate cytidylyltransferase
	Acyl-[acyl-carrier-protein]-UDP-N-acetylglucosamine O-acyltransferase
	Arabinose 5-phosphate isomerase
	Lipid-A-disaccharide synthase
	ADP-L-glycero-D-manno-heptose-6-epimerase
	ADP-heptose synthase

D-glycero-beta-D-manno-heptose 7-phosphate kinase
Transcriptional regulatory protein PhoP
HtrA protease/chaperone protein
Inner membrane protein YrbG, predicted calcium/sodium:proton antiporter
Lipoprotein releasing system transmembrane protein LolE
Lipoprotein-releasing system ATP-binding protein LolD
Lipoprotein releasing system transmembrane protein LolC
Outer membrane lipoprotein carrier protein LolA
Periplasmic chaperone of outer membrane proteins Skp
Outer membrane protein assembly factor YaeT
Proposed peptidoglycan lipid II flippase MurJ
Protein of unknown function YceH
Virulence factor MviM
Beta-1,4-galactosyltransferase
D-alanyl-D-alanine carboxypeptidase
D-alanyl-D-alanine carboxypeptidase
D-alanyl-D-alanine carboxypeptidase
D-alanyl-D-alanine carboxypeptidase
Muramoyltetrapeptide carboxypeptidase
Soluble lytic murein transglycosylase precursor
Multimodular transpeptidase-transglycosylase
Multimodular transpeptidase-transglycosylase
Murein-DD-endopeptidase
Penicillin-insensitive transglycosylase
Phospho-N-acetylmuramoyl-pentapeptide-transferase
UDP-N-acetylglucosamine 1-carboxyvinyltransferase
UDP-N-acetylglucosamine-N-acetylmuramyl-(pentapeptide)
UDP-N-acetylmuramate--L-alanine ligase
Glucose-1-phosphate thymidylyltransferase
Glucose-1-phosphate thymidylyltransferase
dTDP-glucose 4,6-dehydratase
dTDP-glucose 4,6-dehydratase
N-acetylglucosamine-1-phosphate uridyltransferase
N-acetylmannosamine kinase
Phosphotransferase system, N-acetylmuramic acid-specific IIB component
Phosphoglucosamine mutase
Alpha-L-Rha alpha-1,3-L-rhamnosyltransferase
Glycosyl transferase, family 1
Glycosyl transferase, family 2
TRAP-type transport system

TRAP-type C4-dicarboxylate transport system, large permease component
Teichoic acid export ATP-binding protein TagH
YpfJ protein, zinc metalloprotease superfamily
Exported zinc metalloprotease YfgC precursor
Extracellular metalloprotease precursor
Lon protease homolog YcbZ
Rhomboid protease GlpG
Uncharacterized serine protease YdgD
Uncharacterized protease YegQ
Uncharacterized protease YhbU
SOS-response repressor and protease LexA
ClpXP protease specificity-enhancing factor / Stringent starvation protein B
Outer membrane stress sensor protease DegQ, serine protease
Outer membrane stress sensor protease DegS
Intramembrane protease RasP/YluC, implicated in cell division based on FtsL cleavage
HtrA protease/chaperone protein
Protease III precursor
ATP-dependent protease La Type I
ATP-dependent Clp protease ATP-binding subunit ClpX
ATP-dependent Clp protease proteolytic subunit
ATP-dependent Clp protease adaptor protein ClpS
ATP-dependent Clp protease ATP-binding subunit ClpA
ATP-dependent protease subunit HslV
ATP-dependent hsl protease ATP-binding subunit HslU
Intracellular protease
Putative metalloprotease yggG
Catalase KatE-intracellular protease
Tail-specific protease precursor
Protease HtpX
Possible protease sohB
Protease II
Cellulose biosynthesis protein BcsG
Cellulose biosynthesis protein BcsQ
Cellulose biosynthesis protein BcsQ
Cellulose biosynthesis protein BcsE
Cellulose synthase operon protein C
Cellulose synthase operon protein C
Cellulose synthase catalytic subunit [UDP-forming]
Cellulose synthase catalytic subunit [UDP-forming]
Nicotinamide-nucleotide amidase

	N-acetylmuramoyl-L-alanine amidase
	Aliphatic amidase AmiE
	Amidase family protein BBta_1912
	N-acetylmuramoyl-L-alanine amidase
	ADP-ribose pyrophosphatase of COG1058 family
	N-acetylmuramoyl-L-alanine amidase
	Nicotinamidase
	N-acetylmuramoyl-L-alanine amidase
	D-alanine-D-alanine ligase
	D-alanine-D-alanine ligase
	DJ-1/YajL/PfpI superfamily, includes chaperone protein YajL (former ThiJ)
	DJ-1/YajL/PfpI superfamily, includes chaperone protein YajL (former ThiJ)
	Lipopolysaccharide export system protein LptC
	Lipopolysaccharide export system protein LptA
	Lipopolysaccharide ABC transporter, ATP-binding protein LptB
	Lipopolysaccharide biosynthesis protein WzzE
	Lipopolysaccharide N-acetylmannosaminouronosyltransferase
	Lipopolysaccharide core biosynthesis glycosyl transferase kdtX
	Lipopolysaccharide core heptosyltransferase III
	Lipopolysaccharide core heptosyltransferase I
	Lipopolysaccharide assembly protein LapB
	Lipopolysaccharide export system permease protein LptG
	Lipopolysaccharide export system permease protein LptF
	Lipopolysaccharide biosynthesis protein
	Lipopolysaccharide 1,6-galactosyltransferase
	Glycoprotein-polysaccharide metabolism
	Phosphoethanolamine transferase
	Putative periplasmic protein YibQ
	Putative polysaccharide export protein YccZ precursor
	Putative capsular polysaccharide transport protein YegH
	Exopolysaccharide production protein ExoZ
	O-antigen ligase
	O-antigen export system permease protein RfbD
	O-antigen export system, ATP-binding protein
	Uncharacterized protein YhjG
	Virulence factor MviM
Insertion element sequences/transposase	Transposase, IS3/IS911 family
	Transposase, IS3/IS911 family
	Transposase, IS3/IS911 family

	Transposase, IS3/IS911 family
	Transposase, IS3/IS911 family
	Transposase, IS3/IS911 family
	Transposase STM474_p1058
	Transposase STM474_p1058
	Transposase STM474_p1058
	IS3 family transposase
	IS3 family transposase
	ISNCY family transposase
Type III secretion system (T3SS)	Candidate type III effector Hop
Type IV secretion system (T4SS)	TraC
	TraV
Flagellar	Flagellar regulator flk
	Flagellar biosynthesis protein FlhA
	Flagellar biosynthesis protein FlhA
	Flagellar biosynthesis protein FliP
	Flagellar biosynthesis protein FliP
	Flagellar biosynthesis protein FliQ
	Flagellar biosynthesis protein FliQ
	Flagellar biosynthesis protein FlhB
	Flagellar biosynthesis protein FlhB
	Flagellar biosynthesis protein FlgN
	Flagellar basal-body P-ring formation protein FlgA
	Flagellar basal-body rod protein FlgB
	Flagellar basal-body rod protein FlgC
	Flagellar basal-body rod modification protein FlgD
	Flagellar hook protein FlgE
	Flagellar basal-body rod protein FlgF
	Flagellar basal-body rod protein FlgG
	Flagellar L-ring protein FlgH
	Flagellar P-ring protein FlgI
	Flagellar protein FlgJ [peptidoglycan hydrolase]
	Flagellar hook-associated protein FlgK
	Flagellar hook-associated protein FlgL
	Flagellar transcriptional activator FlhD
	Flagellar transcriptional activator FlhC
	Flagellar motor rotation protein MotA
	Flagellar motor rotation protein MotB

	Flagellar protein FlhE
	Flagellar cap protein FliD
	Flagellar biosynthesis protein FliS
	Flagellar biosynthesis protein FliT
	Flagellar hook-basal body complex protein FliE
	Flagellar M-ring protein FliF
	Flagellar motor switch protein FliG
	Flagellar assembly protein FliH
	Flagellar protein FliJ
	Flagellar hook-length control protein FliK
	Flagellar basal body-associated protein FliL
	Flagellar motor switch protein FliM
	Flagellar motor switch protein FliN
	Flagellar biosynthesis protein FliO
	Flagellar biosynthesis protein FliP
	Flagellar biosynthesis protein FliQ
	Flagellar biosynthesis protein FliR
	Flagellar brake protein YcgR
	RNA polymerase sigma factor for flagellar operon
Flagellin	Negative regulator of flagellin synthesis FlgM (anti-sigma28)
	Flagellin protein A
	Flagellin FliC
Type IV pilin, fimbrial and non-fimbrial	NAD(P)H dehydrogenase (quinone), Type IV
	Type IV pilus biogenesis protein PilM
	Type IV pilus biogenesis protein PilQ
	Type IV fimbrial assembly, ATPase PilB
	Type IV pilin PilA
	Type IV fimbrial assembly protein PilC
	Type IV fimbrial assembly, ATPase PilB
	COG3539: P pilus assembly protein, pilin FimA
	COG3539: P pilus assembly protein, pilin FimA
	COG3539: P pilus assembly protein, pilin FimA
	Putative fimbrial biogenesis protein precursor
	Putative pilus chaperone, PapD family
	FIG031703: Fimbriae usher protein StbC
	putative fimbrial usher protein StbD
	Prepilin peptidase dependent protein B precursor
	Prepilin peptidase dependent protein B precursor

	Prepilin peptidase dependent protein A precursor
	Leader peptidase (Prepilin peptidase) (EC 3.4.23.43)/N-methyltransferase (EC 2.1.1.-)
	type 1 fimbriae major subunit FimA
	type 1 fimbriae anchoring protein FimD
	type 1 fimbriae anchoring protein FimD
	type 1 fimbriae anchoring protein FimD
	Fimbriae usher protein StbC
	Outer membrane usher protein fimD precursor
	Sigma-fimbriae usher protein
	Sigma-fimbriae chaperone protein
	Sigma-fimbriae uncharacterized paralogous subunit
	Sigma-fimbriae uncharacterized paralogous subunit
	Sigma-fimbriae tip adhesin
	Sigma-fimbriae usher protein
Phages	Zinc binding domain / DNA primase (EC 2.7.7.-), Phage P4-associated / Replicative helicase RepA, Phage P4-associated
	FIG018226: DNA replication protein, phage-associated
	FIG033266: Phage DNA binding protein
	Phage integrase, Phage P4-associated
	COG3645: Uncharacterized phage-encoded protein
	prophage p2 ogr protein
	Phage baseplate hub
	putative phage repressor
	putative phage repressor
	Phage protein
	FIG107037: Phage late gene regulator
	Phage DNA invertase
Toxin-antitoxin	Ribosome association toxin RatA
	CcdA protein (antitoxin to CcdB)
	CcdB toxin protein
	Succinate dehydrogenase flavin-adding protein, antitoxin of CptAB toxin-antitoxin
	Antitoxin to RelE-like translational repressor toxin
	Death on curing protein, Doc toxin
	Antitoxin DinJ (binds YafQ toxin)
	Antitoxin to RelE-like translational repressor toxin
	Antitoxin to RelE-like translational repressor toxin
	Antitoxin to RelE-like translational repressor toxin
	RelE-like translational repressor toxin
	RelE-like translational repressor toxin
	RelE-like translational repressor toxin

RelE-like translational repressor toxin

RelB/StbD replicon stabilization protein (antitoxin to RelE/StbE)

RelB/StbD replicon stabilization protein (antitoxin to RelE/StbE)

Antitoxin HigA

Antitoxin HigA

Toxin HigB

Toxin HigB

Toxin HigB

Toxin HigB

Toxin HigB

HipB protein @ Antitoxin HigA

Toxin HigB / Protein kinase domain of HipA

VapC toxin protein

Orphan toxin OrtT

Toxic protein SymE

Toxic protein SymE

General discussion

The importance of the study conducted in this thesis is that it showed that bacterial pathogens of onion can reduce seed quality and yield, and can cause serious economic losses. Detection of seedborne pathogens before sowing will help identify potential disease problems and allows steps to be taken to reduce the risks of seed transmission and disease development. Using certified (tested) seed for planting is the starting point for a successful crop as well as an important management tool.

Seed certification is carried out to ensure adequate seed quality and purity. It is a quality assurance process to maintain quality of seeds during seed production, post-harvest and during distribution of seeds. In South Africa, the South African National Seed Organisation (SANSOR) is the chosen body to administer seed certification schemes on behalf of the Ministry of Agriculture, Land Reform and Rural Development. Participation in seed certification is voluntary, but labelling is compulsory. SANSOR follows international [Organisation for Economic Cooperation and Development (OECD) Seed Schemes and the Association of Official Seed Certifying Agencies (AOSCA)] protocols to ensure that seed lots are produced, multiplied and marketed according to standards and systems while maintaining genetic integrity of the final product (seed lots).

Registration of seed companies is required in order to know who is engaged in physical and commercial activities related to seeds, including production, processing, storage, import, export and marketing. Seed certification starts by sourcing healthy, good quality seed. During the registration process, the origin of the seed is verified to ensure that it is acceptable for certification. Field inspections, sampling and quality testing are carried out according to internationally accepted, validated methods and procedures. The fields in which seed companies produce certified seeds are inspected multiple times during the process of planting, production, harvesting, cleaning and packaging. Trained and authorised seed inspectors conduct field inspections, during which aspects such as varietal purity and isolation distances are controlled. The field inspector checks that the crop is satisfactory with regards to varietal identity and purity, freedom from certain weed species, and freedom from key plant pathogens. After harvesting, processing and packaging, the seeds are sampled and tested by registered seed testing laboratories to assess the germination potential and physical purity of the seed lots.

Certified seed lots must comply with the minimum physical requirements as stipulated in the South African Seed Certification Scheme (Table 4 of the Plant Improvement Act (1976) [Appendix A] & Food and Agriculture Organization of the United Nations (FAO), 2006). The sample should be collected by an authorised sampler using the method prescribed by the International Seed Testing Association (ISTA) to ensure that it is representative of the seed lot. The results obtained are presented in a report, issued by the registered seed testing laboratory and submitted to SANSOR (for example) for evaluation and approval before a final seed lot certificate will be issued. The Department of Agriculture, Land Reform and Rural Development (DALRRD) plants post control grow-outs of seed lots to verify the varietal purity and identity of the seed lots.

A seed certification scheme should certify that a container (bag, packet, tin or box) of seed contains what is stated on the label and meets the acceptable standards of seed viability and physical purity, but more specifically the varietal purity and varietal identity of the seed. A container of certified seed has a blue label with the national coat of arms and the words SA CERTIFIED SEED: SANSOR printed on it. Each container is sealed with a SANSOR seal with a unique identification number. In South Africa, regulation 24 of Government Notice Nr R.1064 stipulates that seed of the varieties listed on Table 8 of the Act may only be sold if such seeds have been certified in terms of the SA Seed Certification Scheme.

The seed certification policy enforced by the government encourages the private sector to take part in research and development of new plant varieties. Various schemes are allowed for regulating the quality of seeds produced and distributed, and provide variety protection as per the Seeds Act. This includes the setting up of a National Seed Research and Training Centre to instruct interested individuals in seed technology. Promotion of the informal sectors is an objective of this program to increase the production of seed in order to make the seeds more available, as well as upgrading the quality of farmers' saved seeds; and to develop a national seed network which provides information on accessibility of seed lots of different varieties with production information. Both public and private sector are encouraged to join the network for a clear valuation of demand and supply of seeds. It has been suggested that a national Seed Board be put in place to undertake seed certification and advising of the Government on all matters related to seed planning and development.

As some plant pathogens are able to penetrate and survive within seed, planting pathogen-free seed is an important step towards growing disease-free crops. Onion seed produced in high risk areas should be tested for the following bacteria that can be seedborne in onion: *P. ananatis*, *P. allii*, *P. agglomerans, P. syringae* pv. *allii, P. syringae* pv. *porri and X. axonopodis* pv. *allii*, before sowing. Planting onion varieties in areas where disease pressure is high should be avoided. Seed sanitation procedures can be used to clean the seed, thus eliminating or reducing the incidence of seed with some pathogens. These procedures include hot water treatment and soaking seed in diluted solutions of bleach. The decision on whether or not to use a seed treatment, or which seed treatments to use depends on the condition of the seeds, the plant species, and the particular target pathogens to be controlled.

Summary

Onion (*Allium* cepa L) is widely used in cooking; they add flavour to dishes such as salad, soup and stew. Onions are the third most popular vegetable in South Africa, after potatoes being first and tomatoes being second. South Africa is one of the biggest producer of onion seeds worldwide. Seeds are key input for improving agricultural yield and protecting food security. Providing farmers with quality seeds remains critical to unsure national food security in the country. Seeds quality can be defined in terms of some of the following characteristics: seed vigour, free from seedborne diseases and noxious weeds and should be of proper age. South Africa is one of the biggest producer of onion seeds worldwide. Total volumes of 724.80 tons of onions were produced in the 2014/2015 growing season, of which 657.07 tons were for the export market. Research showed that sales of the top seeds continued to increase, with short day onion seed leading the sales board in South Africa in 2019.

The first chapter of this thesis is a review of the literature, discussing seedborne pathogens of onion focusing on *Pseudomonas* and *Pantoea* species, the mechanisms that these pathogens uses to infect seed. Detection and identification of seed borne pathogens were discussed. Identifying the causal agent of the disease benefits in determining correct approaches to manage or eliminate the pathogen before planting or in the field. Seed treatment aim to promote good seedling establishment, to minimise yield loss, to maintain and improve their quality and to avoid the spread of harmful pathogens. The use of culture-independent methods to examine microbial communities in seeds provides information on diversity of the bacterial communities on and within seed. Plant pathogenic bacteria have developed a number of different mechanisms, which result in disease in the host. Six different secretion systems have been characterised, i.e. T1SS to T6SS. These secretion systems shown to play different roles related to virulence, fitness, colonisation and survival. Full genome sequencing provide information on genetic variations that could lead to disease or can increase the risk of disease development, even in asymptomatic seeds.

For example, in chapter 2, we characterised strains of Pseudomonas syringae, which were isolated from onion plants and seeds. We also included some strains isolated from leek because the strains were isolated from leek plants showing similar symptoms with that observed on onion plants. Biolog GNIII and MLST analysis of four housekeeping genes (*cts*, *gapA*, *gyrB*

and *rpoD*) were used to identify the strains. Both Biolog GNIII dendrogram and MLST analysis showed a cluster supporting the existence of a new pathovar of *P. syringae* and the other corresponding to *P. syringae* pv. *porri*. Pathogenicity of the strains were determined in the glasshouse on onion, leek, chive and spring onion. Pathogenicity results revealed that *P. syringae* pv. *porri* strains induces symptoms on onion, leek, and spring onion. The strains of *P. syringae* of unknown pathovar induced symptoms only on onion. Thus, a new pathovar of *P. syringae* sensu lato, which causes leaf blight of onion, was named pathovar *allii*. The strains of *P. syringae* pv. *allii* differ from the type strain of *P. syringae* CFBP 1392^{T} and *P. syringae* pv. *porri* strains by their ability to produce acid from erythritol and not utilise 3-methyl glucose, D-sorbitol and α-keto butyric acid. This study showed that *cts* (citrate synthase) primers should be used as a quick and accurate means of identifying new strains of *Pseudomonas syringae* to pathovar level.

In this chapter 3, culture-dependent and culture-independent methods were used to evaluate bacterial community of onion seeds in a single cultivar obtained from Northern Cape and Western Cape Province. Culture-independent 16S rRNA-based approach was used because of its ability to detect unculturable bacterial colonisers, as well as those bacteria that are in such low abundance or grow slowly that they could be missed by culture dependent based protocol. The culture-dependent and -independent analyses used in this study indicated that the majority of bacteria associated with fungicide-treated and non-treated onion seed lots were members of the phylum Proteobacteria. The culture-independent approach identified widely recognised plant pathogens or endophytes (e.g. *Pantoea*, *Pseudomonas*, *Streptomyces*, *Stenotrophomonas*; *Sphingomonas*), also some genera that contain species that are potential human pathogens (e.g. *Providencia*, *Enterococcus*, *Sphingobacterium*). Culturable isolates included representatives of *Acinetobacteria*, *Enterobacter*, *Erwinia*, *Microbacterium*, *Pantoea* and *Pseudomonas*. The impact of planting uncertified or untreated seeds will results in loss of yield that will affect the economy.

In chapter 4, whole genome comparative analysis of pathogenic and non-pathogenic *Pantoea agglomerans* was performed to identify genomic differences between pathogenic and non-pathogenic strains in order to reveal possible genetic factors important for emergence of pathogenicity on onion. Genomic analysis revealed that strain BD 1274 of *P. agglomerans*, which is pathogenic on onion, has a larger genome (4, 968 508 bp) than the non-pathogenic strain BD 1212 (4, 875 404 bp), confirming prior observations that non-pathogens lack the

fourth plasmid carrying the T4SS genes responsible for pathogenicity or the four toxin production proteins. Unique genes were identified that differentiated the non-pathogenic and pathogenic *P. agglomerans* strains. Unique genes present in the pathogenic strain only might play a role in bacterial colonisation, fitness, survival and pathogenicity of that strain. The protein encoding the T3SS in the non-pathogenic strain might be important for colonisation and bacterial fitness in onion plants. In contrast, In contrast, conjugal transfer might play a major role in the pathogenicity to onion of strain BD 1274 of *P. agglomerans*. In addition, the toxins identified only in the pathogenic strain could play a significant role in the onion virulence of this strain and the sequences of these toxins can be used for rapid detection of *Pantoea agglomerans*.

Appendix A

Seed standard for onion seed production (Plant Improvement Act (1976) edited in 2018].

TABLE 4 - TABEL 4

PROVISIONS RELATING TO SEED AND SEED SAMPLES - BEPALINGS BETREFFENDE SAAD EN SAADMONSTERS

(*No. of footnote/*No. van voetnota)

Kind of plant		Maximum content (%)			Minimum percentage		Minimum number per weight (kg)	Prepacked seed		Exempted from indication "Prepacked seed"		Maximum mass (kg) of a seed lot
Botanical name	Common name	Other matter	Other seed	Weed seed	Germination by number	Viability	Germination by weighed replicate	Max.mass (g) per container	Approx. no. of seed per container	Max. mass (g) per container	Approx. no. of seed per container	
1	2	3	4	5	6	7	8	9	10	11	12	13
×*Agrotriticum* Cif. & Giacom	Agrotriticum	4	0,2		60			5 000	165 000	500	16 000	20 000
Allium cepa L.	Onion	4	0,2		60			500	170 500	50	17 050	10 000
Allium fistulosum L.	Bunching onion	4	0,2		60			500		50		10 000
Allium fistulosum L. × *A. cepa* L.	New Bunching onion	4	0,2		60			500		50		10 000
Allium porrum L.	Leek	4	0,2		60			500	198 000	50	19 800	10 000
Anthephora pubescens Nees (*7)	Bottle brush grass, (i) Uncoated seed (ii) Coated seed	10 5	0,3 0,3		20 20			5 000 5 000	905 000 145 000	500 500	90 500 14 500	10 000 10 000
Arachis hypogaea L.	Groundnut	2	0,1		70			5 000	5 000 15 000	500	500 1 500	30 000
Asparagus officinalis L.	Asparagus	4	0,2		60			500	25 300	50	2 500	20 000
Avena nuda L. (*6)	Naked oats	0,5	0,3	TR	80			5 000	200 000	500	20 000	30 000
Avena sativa L. [incl. *A. byzantina* K. Koch]	Oats, Red oats	1	0,3	0,1	80			5 000	200 000	500	20 000	30 000
Avena strigosa Schreb.	Black oats	2.5	2,0	0.1	70			5 000		500		30 000
Beta vulgaris L.	Fodder beet, Garden beet, Sugar beet and Swiss chard	6	0,3		60			1 000	58 000	100	5 800	20 000
Brassica napus L. var. *napus*	Oilseed rape	2,0	0,2		70			500		50		10 000
Brassica napus L. var. *napobrassica* (L.) Rchb.	Forage rape and Swede	4	0,2		60			500	157 500	50	15 750	10 000
Brassica oleracea L.	Broccoli, Brussels sprouts, Cabbage, Cauliflower, Fodder kale, Kale, Kohlrabi and Savoy cabbage	4	0,2		60			500	157 500	50	15 750	10 000

Plant Improvement - Table 4