PLANT-ASSOCIATED BACTERIA: BIOLOGICAL, GENOMIC, AND METAGENOMIC STUDIES

By

HEATHER LEE TYLER

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2009

To Mom and Dad

ACKNOWLEDGMENTS

I thank my parents, who have always supported me in my endeavors. I would also like to thank my adviser, Eric W. Triplett. Lastly, I would like to thank all the members of the Triplett Lab, both past and present, for their advice and support. Without them, I would not have survived the rough times when nothing seemed to work and I thought I'd never make it through.

TABLE OF CONTENTS

page

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

PLANT-ASSOCIATED BACTERIA: BIOLOGICAL, GENOMIC, AND METAGENOMIC STUDIES

By

Heather Lee Tyler

December 2009

Chair: Eric W. Triplett
Major: Microbiology and Cell Science

Bacteria can interact with plants in a variety of ways, as pathogens causing disease, as epiphytes living on plant surfaces, or they can live within plant tissue without causing disease as symbionts or endophytes. Therefore, a range of different plant-bacterial interactions was studied in this work. A metagenomic study utilizing high-throughput sequencing data from citrus tissue with the disease, Huanglongbing (HLB), was conducted to assess the microbial community in infected tissue. The only bacterium identified in infected tissue was the proposed causative agent of HLB, '*Candidatus* Liberibacter asiaticus.' As this bacterium has yet to be cultured, this work substantiates its proposed role in disease development. Plant growth-promoting endophytes were also examined, looking at the genomic level as well as at specific plant growth-promoting responses. Using optical mapping, two contradictory genome sequences of *Gluconacetobacter diazotrophicus* PAl 5 were distinguished, identifying which was the best representation of the actual PAl 5 chromosome. The results of this study are important since the accuracy of the genome sequence could have significant impacts on comparative genomic analyses with this strain. The plant growth-promoting mechanisms used by two endophytic enteric bacteria, *Klebsiella pneumoniae* 342 and *Enterobacter cloacae* P101, were also investigated, utilizing their genome sequences to guide experiments. Both bacteria increased

lateral root numbers on *Arabidopsis* plants and genome annotation indicated each posses the auxin synthesis gene, *ipdC*. As auxin is known to play a role in lateral root development, subsequent work on lateral root promotion focused on this phytohormone. As part of this work, a plasmid was constructed that expressed the *ipdC* gene and could be stably maintained without selection pressure. While these enteric endophytes are related to human pathogens and Kp342 was found to have pathogenic potential, the information gained by studying these bacteria could yield insight into enhancing the growth-promoting effects of other endophytes. Thus, the work described here exemplifies how metagenomic, genomic, and biological experiments can be used to gain both an overview and in depth view of plant-bacterial associations.

CHAPTER 1
INTRODUCTION

Bacteria can interact with plants in a variety of different ways. Many plant-associated bacteria are pathogens that cause disease in plants. One plant disease studied in this work is known as Huanglongbing (HLB), and the bacterium proposed to cause this disease, '*Candidatus* Liberibacter spp.', has yet to be cultured (Bové 2006). There are also many other bacteria that live in association with plants without causing disease symptoms. These bacteria can be epiphytes living on plant surfaces, symbionts living within plant tissue and creating symbiotic structures, or endophytes living in plant tissue without forming symbiotic structures. Such growth-promoting bacteria examined in this work include *Gluconacetobacter diazotrophicus* PAl 5, *Klebsiella pneumoniae* 342, and *Enterobacter cloacae* P101 (Chelius and Triplett 2000; Riggs et al. 2001; Sevilla et al. 2001).

Alarmingly, some plant-associated bacteria are also related to human enteric pathogens. Colonization of plants with such bacteria can lead to disease when they inhabit produce crops that are eaten raw. As a result, considerable effort has gone into studying how these human pathogens colonize and interact with plants. The ultimate goal of such research has been to gain enough insight into the colonization of plants by enteric human pathogens in order to limit it, thus making food supplies safe for human consumption.

Alternatively, plant-associated bacteria are also studied extensively due to their positive effects on plant growth. In contrast to the work done on human pathogens, these plant growth-promoting strains have been studied in the hopes of enhancing the interaction rather than limiting it. In this way, knowledge of plant growth-promoting bacteria can be extended toward increasing production in plant crops. Interestingly, some of the plant-associated bacteria found to enhance plant growth turn out to be related to human pathogens, as was the case with *K.*

pneumoniae 342 used in the current study. As a result, it is important to check even plant growth-promoting isolates for pathogenic potential before considering their use in agricultural settings.

Regardless of whether plant growth-promoting bacteria are human pathogens, study of such bacteria can still provide valuable information. For example, the genome sequences of these bacteria can be used in comparative analyses with the genomes of non-plant associated strains as well as the genomes of other plant-associated bacteria to identify genes unique to plant colonization and growth promotion. Such genes could be engineered into nonpathogenic bacteria to enhance their colonization and growth-promoting effects on plants. Therefore, even if a bacterial species cannot be used in agricultural settings due to pathogenicity, knowledge of the plant colonization and growth-promoting mechanisms these bacteria utilize may be applied to enhance other, nonpathogenic plant growth-promoting bacteria.

Given the breadth of interactions seen between plants and bacteria, and the increasing emphasis on genomic analysis of plant-associated bacteria, a range of interactions between plant-associated bacteria and their hosts were studied in this work using a variety of different approaches. The literature on plant-associated bacteria was reviewed and has been published in Annual Review of Phytopathology (Tyler and Triplett 2008). Though the emphasis of this literature review was on the presence of human pathogens in plants, plant growth-promoting bacteria and the prospects of using genomics in the study of these plant-associated bacteria are also discussed. In regards to examining plant pathogens, research using metagenomics to identify bacteria within citrus tissue infected with the disease, HLB, is also presented. This work has been accepted for publication at Molecular Plant-Microbe Interactions (Tyler et al. in press), and is significant not only for the knowledge it contributes to the study of this devastating citrus

disease, but also as an example of how metagenomics can be utilized to study other uncultured plant pathogenic bacteria.

When examining plant growth-promoting bacteria, work was also performed to distinguish between contradictory genome sequences of *G. diazotrophicus* PAl 5 and has been submitted for publication to BMC Genomics. Such work is important as the accuracy of genomic sequences could significantly impact the results of comparative genomic analyses planned for plant growth-promoting bacteria in the future. Lastly, while genomic and metagenomic-based studies are useful in providing global views of plant-bacterial interactions, it is still important to examine specific mechanisms in these interactions to confirm the observations seen in these studies, for sequence analysis can only provide so much information without verification in the actual bacterium. Therefore, the plant growth-promoting effects of *K. pneumoniae* 342 and *E. cloacae* P101 were also examined using a variety of assays, ultimately utilizing genomic sequences to guide experiments on the growth-promoting phenotype. Thus, the work presented here aims to examine several aspects of plant-bacterial interactions, ranging in depth from examining entire genomes to studying specific mechanisms utilized by plant growth-promoting bacteria.

CHAPTER 2
PLANTS AS A HABITAT FOR BENEFICIAL AND/OR HUMAN PATHOGENIC BACTERIA

Introduction

Non-plant pathogenic bacteria are common inhabitants of the interior of plants. These bacteria can promote plant growth, improve the nitrogen nutrition of plants, and be pathogenic to humans. Early work in bacterial endophytes was encouraged by their ability to improve plant growth. As endophytic bacteria were discovered in a variety of host plants, it became clear that enteric bacteria were frequent inhabitants of the plant cortex. These observations coupled with some well-publicized disease outbreaks from the consumption of raw alfalfa sprouts in the late 1990s led a number of groups to study the ability of enteric pathogens to colonize the interior of plants.

Human pathogens in raw produce took on national prominence in the United States between August and October of 2006 when 205 illnesses in 26 states were reported as a consequence of the consumption of spinach contaminated with *Escherichia coli* O157:H7. This outbreak resulted in 105 hospitalizations, 31 hemolytic uremic syndrome cases, and three deaths (Calvin 2007). The geographic scale of this epidemic startled the public and fostered a renewed interest in research in this area. However, these outbreaks are not uncommon. According to the Centers for Disease Control and Prevention's data (Lynch et al. 2006), from 1998 to 2002, an average of 25,674 food borne illnesses and 18 deaths in the United States were caused by 1329 outbreaks annually. The number of outbreaks traced to plant consumption (vegetables, fruits, and nuts) during that same period averaged 56 and caused 2,109 illnesses each year. Although the proportion of outbreaks attributed to plant consumption is just 4.2%, outbreaks from plants appear to be more serious for human health as they cause 8.2% of the illnesses and 8% of the deaths. Assuming that 750 million meals are consumed each day in the United States

(roughly 2.5 meals per day per person), food borne illnesses caused by plant consumption are very rare and our food supply is remarkably safe.

Nevertheless, these outbreaks are well publicized and cause considerable alarm. Any outbreak stemming from raw produce can cause far more harm if not reported quickly and followed by rapid source identification. Plant-derived outbreaks can be caused by many factors: Pathogens may be located on the plant surface or within plant tissues; pathogens may be unwittingly added to raw produce during food preparation.

In this review, we discuss one source of these pathogens in plants. What is known about the ability of enteric bacteria to enter plants? To what extent do plant hosts regulate their entry? How do clinical enteric pathogens differ from plant-derived enteric pathogens? What do we still need to know to address this problem? How can improved knowledge lead to a reduction in outbreaks? For more information about the epidemiology and fitness of human pathogens on plant surfaces, the readers are referred to the recent review by Brandl (2006). For a review of the diversity of bacteria in plants, the readers are referred to Rosenblueth and Martinez-Romera (2006).

Cycle of Human Pathogens in the Environment

Human pathogens have been found in plants in the environment as well as in crop plants. In nature, a high number of facultative human pathogens have been associated with the bryophyte *Sphagnum*, including *Staphylococcus*, *Hafnia*, *Yersinia*, and *Pantoea* (Opelt et al. 2007). In addition, *Pantoea*, *Salmonella*, *Enterobacter*, *Citrobacter*, and *Klebsiella* have been found to endophytically colonize the leguminous tree, *Conzattia multiflora* in Mexico (Wang et al. 2006). The presence of *Salmonella* in this plant indicates that plants in natural settings may harbor human pathogens (Wang et al. 2006). Several members of the Enterobacteriaceae have been isolated from both wild and cultivated plants, including dandelions, plantains, tomatoes,

and potatoes (Markova et al. 2005). Though none of the species isolated in this work were known pathogens, many were resistant to multiple antibiotics and capable of adhering to human erythrocytes (Markova et al. 2005).

More widely known and studied is the occurrence of human pathogens in crop plants due to fertilizing and watering with contaminated waste materials. High incidences of microbial contamination, including species of *E. coli*, *Enterobacter cloacae*, and *Klebsiella pneumoniae*, have been reported in several vegetable species irrigated with untreated wastewater (Ibenyassine et al. 2007). Other cases of potential pathogens in crops include the isolation of *Listeria monocytogenes* and *E. coli* from organically grown lettuce and *Enterobacter* spp., *Klebsiella*, and *E. coli* on mung bean, though the enterohemorragic *E. coli* O157:H7 strain was not identified in either of these cases (Loncarevic et al. 2005; Robertson et al. 2002).

One explanation for the occurrence of human pathogens in plants is that they cycle through the environment, using plants as an alternative host to survive in the environment and as a vehicle to recolonize animal hosts once ingested. To accomplish this process, enteric pathogens shed in animal feces must be transported to and come in contact with plants eaten by the host. One possible mechanism for this transfer is through nematodes, such as *Caenorhabditis elegans*, which are capable of transporting *Salmonella newport* to fruits and vegetables through soil (Kenney et al. 2006). When soil inoculated with *C. elegans* was placed on top of contaminated manure, the bacteria were found on produce above the soil surface, whereas the bacteria were generally not found on the vegetables when *C. elegans* was absent from soil (Kenney et al. 2006).

Other possible mechanisms of pathogen transport to plants include irrigation water and runoff from livestock pastures. Numerous instances have been reported whereby contaminated

irrigation water led to contaminated crops. In a field study, a wide array of produce plants, including lettuce, parsley, carrots, and tomato, had elevated concentrations of fecal indicator organisms when irrigated with highly polluted effluents, whereas irrigation with only slightly polluted water resulted in lower numbers of indicator organisms (Armon et al. 1994). In another study, the same *Salmonella* strains present in edible parts of vegetables were found in the contaminated raw wastewater used to irrigate the plants (Melloul et al. 2001). Vegetables that grow on the surface of the ground, such as lettuce or parsley, were more contaminated than those that develop above the surface, such as tomato (Melloul et al. 2001). *Salmonella* can also contaminate alfalfa sprouts produced with contaminated water. A strong correlation was observed between the numbers of *Salmonella enterica* cells present in alfalfa sprouts and in the irrigation wastewater (Howard and Hutcheson 2003). The population of *S. enterica* cells was able to grow in the wastewater, using compounds in it as a sole carbon and nitrogen source (Howard and Hutcheson 2003). In addition, the bacteria multiplied rapidly on alfalfa during the first 24 hours of the germination process (Howard and Hutcheson 2003). The ability to grow on alfalfa was independent of serovar, source of isolation, or virulence, with isolates from meat and stool reaching densities similar to those of plant isolates (Howard and Hutcheson 2003). This rapid establishment on alfalfa sprouts and similar colonization by all strains tested demonstrates how *Salmonella* could pose a significant problem in contaminated plants meant for human consumption.

In addition to culturing bacteria, polymerase chain reaction (PCR) of repetitive deoxyribonucleic acid (DNA) sequences has also been used to study the role of irrigation water in contaminating plants. Enteropathogenic *E. coli* strains present in irrigation water were also detected on vegetables and in soil by this method, demonstrating that bacteria can be transported

by irrigation water and result in contamination of produce (Ibenyassine et al. 2006). The irrigation method employed also appears to play a role in contamination: Greater contamination of plants was found with furrow irrigation than with subsurface drip irrigation (Song et al. 2006).

Human pathogens contaminate crops from both contaminated manure and irrigation water. When exposed through either contaminated manure or irrigation water, *E. coli* O157:H7 persisted in soil for 154 to196 days after treatment and were detected on onion and carrots for 74 and 168 days, respectively (Islam et al. 2005). In the case of *S. enterica* serovar Typhimurium, bacteria survived in soil 203 to 231 days and were found on radishes and carrots 84 and 203 days after sowing in soil with contaminated manure or irrigation water (Islam et al. 2004). Thus, both contaminated soil and tainted irrigation water can contribute to colonization of crops by human pathogens.

Another mechanism for the transfer of pathogens is runoff from contaminated areas such as cow pastures that spreads bacteria to land where crops are cultivated. For example, isolates of *E. coli* can be transported rapidly in runoff, with attachment to soil particles affecting how far they travel (Muirhead et al. 2006). When not preattached to larger particles (>45 μm), *E. coli* in runoff attach to much smaller fragments (<2 μm) and are not slowed or deposited during runoff transport (Muirhead et al. 2006). Cattle and other livestock animals kept in close proximity to croplands are a potential source for pathogenic bacteria in runoff and manure. Ruminant diet appears to play a role in survival of certain pathogens in cow feces: *E. coli* O157:H7 lived longer in manure from cows eating grass and maize silage than in manure from cows on a pure straw diet (Franz et al. 2005). In addition, the pathogen *Campylobacter jejuni* was found in the feces of feedlot cattle (Besser et al. 2005). The presence of these bacteria in the feces of

livestock could possibly lead to their spread to crop plants through runoff or from the use of manure from these cattle as soil amendments.

Another major factor important to colonization of plants is how long these bacteria persist in the soil before dying. *E. coli* O157:H7 present in animal waste was capable of surviving in soil, although the source of the waste did have an effect on survival (Williams et al. 2007). Bacteria from ovine stomach content waste survived in higher numbers than bacteria from cattle slurry (Williams et al. 2007). *E. coli* O157:H7 was also found to persist in soil for 8 weeks after treatment with contaminated manure, though it was not found in or on lettuce leaves or roots (Johannessen et al. 2005). The lack of *E. coli* O157:H7 on any part of lettuce may be due to the presence of *Pseudomonas flourescens* in the rhizosphere, which was found to inhibit O157:H7 in culture (Johannessen et al. 2005). The use of Ecosan sludge in agriculture has also been linked to bacterial contamination of produce. The application of sludge containing *Salmonella*, fecal coliforms, and fecal streptococci to soil increased the bacterial counts in spinach and carrot plants, with increasing amounts of sludge resulting in higher bacterial numbers in plants (Jimenez et al. 2006). *Listeria innocua* and *Clostridium sporogenes* present in contaminated soil amendments have also been found to persist in the soil and were both found on the leaves of parsley grown in the contaminated soil, though they were not observed within the plants (Girardin et al. 2005). While *Clostridium* spores persisted in soil for 16 months, *Listeria* decreased by 10,000,000-fold over 90 days (Girardin et al. 2005).

Because plants can serve as hosts for enteric human pathogens in the environment, these enterics may be adapted to surviving longer and in higher numbers in soil when the plants are present. Gagliardi and Karns (2002) found that the presence of certain crops increased the persistence of *E. coli* O157:H7 in soil. In unplanted, fallow soils, *E. coli* O157:H7 persisted only

for 25 to 41 days, but was found up to 92 and 96 days if alfalfa or rye plants were present, respectively (Gagliardi and Karns 2002). Ibekwe and associates (2004) also found that the presence of roots in contaminated soils increased concentrations of *E. coli* O157:H7. In this study, *E. coli* introduced through irrigation water was found to reach higher densities in rhizosphere soils than in non-rhizosphere soils (Ibekwe et al. 2004). In both studies, clay was found to increase the persistence of *E. coli* O157:H7 and other coliforms (Gagliardi and Karns 2002; Ibekwe et al. 2004). In addition, bacterial populations in soil may increase after the addition of plants to soil. For example, the bacterial population in soil spiked after the addition of plant material and then fluctuated in a wave-like fashion that was not found to be the result of nitrogen shortages or pH (Zelenev et al. 2005). In contrast to these results, the presence of maize roots did not effect survival of *E. coli* in soil (Williams et al. 2007). Likewise, the presence of legume crops other than alfalfa did not increase persistence of *E. coli* more than in fallow soils (Gagliardi and Karns 2002).

In addition to direct colonization from exposure to contaminated soil and water, some pathogens that establish themselves endophytically in plants may be seedborne and inherited from generation to generation. In *Conzattia* trees, nonpathogenic bacteria that were isolated as endophytes were also found in all seeds that were tested, indicating that these endophytes may be seedborne and inherited from the previous generation (Wang et al. 2006). In *S. enterica* and *E. coli* O157:H7 on *Arabidopsis*, the bacteria were occasionally found to contaminate seeds, and this contamination was highly correlated with contaminated chaff (Cooley et al. 2003). *Salmonella* survived both in and on tomato throughout plant growth, from inoculation to flowering and fruit ripening (Guo et al. 2001). *Salmonella* were found in tomato fruit from plants inoculated on the stem both before and after flower set, with 43% and 40% of tomatoes

from these treatments testing positive for the bacteria, respectively (Guo et al. 2001). Although most inoculated flowers were reported to abort, 25% of the tomatoes from inoculated flowers harbored *Salmonella* (Guo et al. 2001). In apples, *E. coli* O157:H7 were observed attached to seed integuments, with infiltration occurring through the blossom's calyx and traveling up the floral tube to the internal parts of the apple (Burnett et al. 2000). These observations support the hypothesis that the enteric pathogens colonizing produce can spread to the seeds and be transmitted to the next generation.

Observations of Pathogens on Plant Surfaces

Strains of enteric pathogens differ in their ability to colonize the surface of plants. For example, cells of *E. coli* O157:H7 were rinsed off after repeated washes, whereas *E. coli* serotypes isolated from cabbage roots that had been exposed to sewage adhered to sprouts (Barak et al. 2002). In contrast, all *S. enterica* serovars tested did not differ among themselves in binding to alfalfa, and they all bound significantly better than did *E. coli* O157:H7 (Barak et al. 2002). Surface moieties on cells are not needed for the initial attachment of *E. coli* O157:H7 to lettuce, since nonbiological FluoSpheres, as well as live and dead *E. coli* cells were retained on plants at similar densities (Solomon and Matthews 2006). This, and the observation that *E. coli* did not bind as strongly to plants as *Salmonella*, demonstrate that not all pathogens are as disposed to colonize plants as others. Given that strains vary in their ability to bind to plant tissue (Barak et al. 2002), the *E. coli* strain used in this study may be one that is not efficient at attachment.

E. coli did not grow on plants as well as *Salmonella*, with *E. coli* O157:H7 growing significantly less than *S. enterica* on sprouting alfalfa (Charkowski et al. 2002). *Salmonella* also colonized both the seed coat and sprout roots, whereas O157:H7 only colonized alfalfa roots (Charkowski et al. 2002). One commonality between colonization by these two bacteria is that

both temperature and inoculum density affected initial colonization by both strains, with both higher temperatures and greater inoculum densities resulting in higher colonization (Charkowski et al. 2002). Temperature is also important to the colonization of the cilantro phyllosphere by *S. enterica* serovar Thompson, which was isolated from a cilantro-related outbreak (Brandl and Mandrell 2002). The pathogen did not colonize as well as other plant surface bacteria at 22°C, but reached higher numbers when grown at higher temperatures, most likely by outcompeting the other epiphytes due to its higher growth rate (Brandl and Mandrell 2002).

Endophytic Colonization

Endophytic colonization of human pathogens into plants has been observed in several studies, but how the bacteria enter these plants and where they are localized within plants are not completely understood. The internalization of human pathogens into produce is of particular interest since it may protect these bacteria from sanitizing treatments meant to make fruits and vegetables safe for consumption. Additionally, the amount of endophytic colonization varies according to the bacterium and plant species examined. From numerous observations of bacterial colonization of plant roots, it is widely assumed that bacteria can colonize and enter roots at sites of lateral root emergence. Both *Salmonella* and *E. coli* have been observed by confocal microscopy to invade plant roots through lateral root cracks of *Arabidopsis thaliana* (Cooley et al. 2003). *E. coli* O157:H7 was also observed colonizing preferentially at root junctions on lettuce (Jablasone et al. 2005). Confocal microscopy of the endophyte *K. pneumoniae* 342 on plants revealed significant colonization around lateral root cracks of *Medicago sativa*, *M. truncatula*, *A. thaliana*, *Triticum aestivum*, and *Oryza sativa*, suggesting a possible entry site for the bacteria into the plant (Dong et al. 2003a and b). In addition, *Serratia* cells were seen at higher concentrations around lateral root emergence sites of rice, again

indicating that these sites may be potential entry points for bacteria into plants (Gyaneshwar et al. 2001).

Several studies have focused on internalization of the enteric pathogen *E. coli* O157:H7, related species of *Salmonella*, and *Listeria*. *E. coli* has been observed to be internalized through the root system of maize grown in contaminated hydroponic media and was found in the shoot of plants 48 hours after inoculation (Bernstein et al. 2007). Damaging the root system by cutting off root tips or removing roots at the root shoot junction increased the number of *E. coli* internalized in shoots 27.8 and 23.9 times greater than in undamaged roots (Bernstein et al. 2007). Though capable of colonizing maize roots, *E. coli* was observed to decline over time, but a sharp reduction in population was not observed until 4 days post inoculation, and bacteria were still present at approximately 10^2 colony forming units (CFU)/g fresh weight after 7 days (Bernstein et al. 2007). *E. coli* O157:H7 has also been found to internalize within radish sprouts and carrots (Auty et al. 2005; Itoh et al. 1998). *E. coli* found in root-inoculated radish sprouts was observed within stomata of cotyledons and the inner tissues of edible hypocotyls and cotyledons (Itoh et al. 1998). In carrots, *E. coli* was found at cell junctions and intercellular spaces up to 50 μm within carrot tissue, but did not penetrate the carrot cells (Auty et al. 2005).

Several studies have observed the internalization of *Escherichia coli* O157:H7 within lettuce plants (Franz et al. 2007; Solomon et al. 2002; Wachtel et al. 2002). *E. coli* O157:H7 has been found within the tissue of thoroughly surface-sterilized lettuce leaves at a density of 3.95 log CFU/g (Franz et al. 2007). Two studies found that *E. coli* O157:H7 can be internalized into lettuce when exposed to contaminated irrigation water and manure-treated soil (Solomon et al. 2002; Wachtel et al. 2002). Wachtel and associates (2002) observed *E. coli* O157:H7 binding preferentially to the roots of lettuce from contaminated irrigation water, and also observed strains

in the edible portion of the plant. In confirmation of this observation, fluorescently labeled bacteria could be seen moving in the vasculature of a hypocotyl, most likely in the xylem (Wachtel et al. 2002). When grown in soil, *E. coli* O157:H7 was associated with the plants 3 days after inoculation, even with low inoculum doses, indicating that the bacteria could multiply in the plant (Wachtel et al. 2002). Solomon and associates (2002) observed *E. coli* O157:H7 up to a depth of 45 μm within lettuce leaf tissue in aggregates in intercellular spaces and found that direct contact with contaminated material was not required for internalization of the bacteria within edible leafy parts of the plant.

Strains of *Salmonella* have also been observed within lettuce. *S. enterica* serovar Typhimurium was found in high numbers of 2.57 log/g in surface-sterilized lettuce (Franz et al. 2007). The ability to internalize and endophytically colonize lettuce is not common to all *Salmonella* strains. In a study of three *S. enterica* serovars, all strains were capable of colonizing the plant epiphytically, but only the Dublin serovar could colonize endophytically (Klerks et al. 2007). Differences between *Salmonella* strains have also been observed in tomato, with Montevideo being the most persistent and Poona the most dominant (Guo et al. 2001). *S. enterica* serovar Typhimurium colonizes barley (*Hordeum vulgare*) roots endophytically up to 10^6 CFU/g fresh weight (Kutter et al. 2006). Fluorescence in situ hybridization and confocal laser microscopy confirmed endophytic colonization, and revealed that *Salmonella* had colonized root cells and was present in the inner cortex of the roots (Kutter et al. 2006).

Endophytic colonization by human pathogens has been studied in several other plants. In *A. thaliana*, *S. enterica* and *E. coli* O157:H7 colonize the entire plant following root inoculation, indicating that the bacteria can migrate on or in plants and are detectable in the plants for up to 21 days after inoculation (Cooley et al. 2003). Internalization of enteric pathogens has also been

investigated in legume species. Both *E. coli* and *Salmonella* Montevideo were found to internalize into mung bean sprouts from inoculated seeds (Warriner et al. 2003). These bacteria were observed in roots and between epidermal cells on hypocotyls and were isolated from surface-sterilized tissue and apoplastic fluid, indicating the bacteria were internalized and protected from sterilization treatments (Warriner et al. 2003). Also, *E. coli* was visualized in the vascular system of mung bean sprout hypocotyls using β-glucuronidase (GUS) staining (Warriner et al. 2003). In testing several pathogenic strains of *E. coli* and *S. enterica* for endophytic colonization of alfalfa (*M. sativa*) and its relative, *M. truncatula*, Dong and associates (2003a) found that the strains varied greatly in their ability to colonize plants. Although endophytic colonization was observed by all strains tested, significant variability in the extent of endophytic colonization between strains was observed (Dong et al. 2003a). The pathogenic strains were compared to a plant-isolated endophyte, *Klebsiella pneumoniae* 342 and a lab strain of *E. coli* K12 and were found to have colonized the plants at a level between these two strains (Dong et al. 2003a).

Endophytic colonization of *E. coli* O157:H7 within fruits has also been studied. In apple, scanning confocal laser microscopy (SCLM) has been used to observe *E. coli* O157:H7 attached preferentially to discontinuities in the cuticle of apples and at puncture wounds down to 70 μm in the tissue (Burnett et al. 2000). *E. coli* O157:H7 has also been observed colonizing apple lenticels, in some cases to a depth of 40 μm (Burnett et al. 2000). Bruising and rubbing of apples appears to increase the depth to which bacteria can internalize within lenticels (Kenney et al. 2001). Beyond infiltration from the surface, *E. coli* O157:H7 can colonize the interior of apples in the core and subsurface structures, attaching to the cartilaginous pericarp and seed integuments with infiltration occurring through the blossom's calyx and traveling up the floral

tube to the internal parts of the apple (Burnett et al. 2000). In wounded apples, *E. coli* O157:H7 was observed to colonize and survive by causing cell membranes to degrade and release their contents (Janes et al. 2005). *E. coli* within apple tissue formed granules and vesicles that were not present in bacterial growth in brain heart infusion broth, possibly due to osmotic conditions in the apple tissue (Janes et al. 2005).

In oranges, *E. coli* O157:H7 and *Salmonella* colonize the interior of fruit at the same rate as the entry of a dye, suggesting a passive mechanism of entry (Eblen et al. 2004). Oranges with large puncture wounds were more likely than ones with smaller wounds to internalize the bacteria (Eblen et al. 2004). Both *E. coli* and *Salmonella* were able to grow in oranges at 24°C (Eblen et al. 2004). The roots and xylem of citrus plants can also be colonized by enteric bacteria. *K. pneumoniae* 342 (Kp342) was able to endophytically colonize both *Citrus sinensis* and *Catharanthus roseus*, though it was present in higher numbers in *C. roseus* plants (Lacava et al. 2007). The bacteria that were inoculated onto roots were subsequently found in the xylem of roots and branches of both species tested, with Kp342 colonies observed in the xylem by fluorescent microscopy of seedling cross-sections (Lacava et al. 2007).

Kp342 has also been found to colonize the interiors of *M. sativa* (alfalfa), *M. truncatula*, *A. thaliana*, *T. aestivum* (wheat), and *O. sativa* (rice) (Dong et al. 2003b). Kp342 was more efficient at colonizing the plant apoplast compared to the type strain of *K. pneumoniae* (Dong et al. 2003b). Differences in colonization levels between plants were also observed, with monocots being colonized in higher numbers than dicots (Dong et al. 2003b). Enteric bacteria isolated from plant tissues colonized the interior of plants in higher numbers than those isolated from clinical or other settings. For example, of several *Salmonella* strains tested for endophytic colonization of alfalfa seedlings, strain SCH7976 colonized the interior more than the other

strains. This strain was derived from an alfalfa sprout salmonellosis outbreak in California in 1998. Kp342, which was originally isolated from a nitrogen-efficient line of maize (Chelius and Triplett 2000), colonizes plants in far higher numbers than any other enteric bacterium tested to date (Dong et al. 2003a and b).

Damage to plant tissue can provide points of entry for bacteria into plant tissue. For example, bacteria were not detected on the surface of unbroken spinach leaves using scanning electron microscopy, but they were present on those leaves where the cuticle was broken, spreading to the internal palisade parenchyma (Babic et al. 1996). In this case, strains of the Pseudomonadaceae, Enterobacteriaceae (including *K. pneumoniae*), and *Micrococcaceae*, and lactic acid bacteria were found (Babic et al. 1996).

Bacterial Competition

When colonizing plants, human pathogens undoubtedly come into contact with native bacteria and must overcome microbial competition to establish themselves. These interactions vary according to which pathogens and host plants are being studied and which bacteria are used as antagonists. When looking at *Salmonella* growth on alfalfa sprouts, coinoculation with bacteria isolated from market sprouts reduced *Salmonella* numbers more than coinoculation with *P. fluorescens* 2-79 seven days after inoculation, whereas microbial communities from lab-grown sprouts showed the least reduction on *Salmonella* growth (Matos and Garland 2005). The observation that microbial communities from market sprouts reduced *Salmonella* numbers more than those from lab-grown sprouts suggests that bacteria undergo selective conditions in the field and/or during handling on the way to market that result in a more robust native community for competition against contaminating pathogens (Matos and Garland 2005). This conclusion is supported by the observation that market sprout microbial communities could grow faster on a wide array of substrates (Matos and Garland 2005).

In most cases, rhizobacteria did not reduce the persistence or number of enteric pathogens on plants, with the exception of *E. cloacae*, which reduced the numbers of *E. coli* O157:H7 and *L. monocytogenes* on lettuce (Jablasone et al. 2005). Another *Enterobacter*, *E. asburiae*, was also observed to be antagonistic toward *E. coli* O157:H7 on plants, lowering colonization on lettuce by 20- to 30-fold (Cooley et al. 2006). This effect is most likely due to competition for carbon and nitrogen sources, since *E. asburiae* and *E. coli* both utilize over 20 of the carbon sources present in plant exudates (Cooley et al. 2006). The conditions on the plant where bacteria initially land may also play a role in their survival. Immigrant bacteria of both *P. agglomerans* and *P. flourescens* were twice as likely to survive on leaf surfaces when they landed on aggregates of previously established bacteria than when they landed on uncolonized areas (Monier and Lindow 2005). In addition, the plant structures on which these bacteria landed also influenced their survival, with bacteria landing on veins and granular trichomes having higher survival rates than those landing on hooked trichomes (Monier and Lindow 2005). In addition, bacteria landing on epidermal cells altered by *P. agglomerans* (which were not necessarily coincident) had higher survival rates than if they had landed on normal, unaltered epidermal cells (Monier and Lindow 2005). It would be interesting to see if similar relationships are seen in human pathogens colonizing plant surfaces.

Role in Plant Growth Promotion

Many plant-associated bacteria are known to have beneficial effects on plant growth, and the capability of some animal pathogens to enhance plant growth may indicate an evolutionary reason why these pathogens evolved to colonize plant hosts, with the plant benefiting from the bacteria, and the bacteria gaining nutrients from the plant and protection from the environment. There are numerous instances in the literature of bacterial plant growth promotion, so this review gives only a brief summary of the topic as it relates to potential human pathogens. Several

genera of endophytes isolated from the leguminous tree, *C. multiflora*, were identified as members of the Enterobacteriaceae family, including *Enterobacter*, *Klebsiella*, *Citrobacter*, and *Pantoea*, and were shown to increase the height of seedlings after inoculation (Wang et al. 2006). Plant-associated bacteria can increase plant growth by several mechanisms, the most commonly studied of which involve production or lowering of plant hormones important to development. Several plant hormones have been identified in the culture supernatants of *K. pneumoniae*, *E. coli*, *Proteus vulgaris*, and *Bacillus cereus*, including auxin, gibberellin, and cytokinin, with *Klebsiella* producing significantly more auxin than the other strains (Karadeniz et al. 2006). El-Khawas and Adachi (1999) also identified auxin in *K. pneumoniae* supernatants and found that application of the supernatants to rice roots resulted in increased root length, surface area, and fresh weight. The role of bacterial auxin production has been further confirmed in studies of other plant-associated bacteria. Knocking out the *ipdC* auxin synthesis gene in *Azospirillum brasilense* lowered auxin production by 90% and resulted in reduction of the growth-promoting effect this bacteria has on the root system (Barbieri and Galli 1993; Dobbelaere et al. 1999). Also, in *Pseudomonas putida* GR12-2, the wild type strain increased root length 35% to 50% more than an *ipdC* mutant in mung bean (Patten and Glick 2002).

In addition to secreting plant hormones, some plant growth-promoting bacteria produce enzymes that interfere with and break down hormones that inhibit plant growth. For example, 1-aminocyclopropane-1-carboxylate (ACC) deaminase, an enzyme that lowers the concentration of ethylene in plants by cleaving ACC, the precursor of this hormone, is found in several plant growth-promoting bacteria (Glick 2005). Inoculation with an ACC deaminase producing strain, *E. cloacae* UW4, resulted in increased root length on canola seedlings, whereas an *acdS* mutant in this strain that does not possess ACC deaminase activity is diminished in its ability to enhance

root length (Li et al. 2000). In addition, transferring the ACC deaminase gene from *E. cloacae* UW4 into a *P. fluorescens* strain lacking this enzyme confers the ability to increase root length in canola, suggesting that bacterial ACC deaminase lowers ethylene concentrations inside the plant, resulting in an increase in root length (Wang et al. 2000). ACC deaminase's ability to enhance root growth has also been linked to changes in gene expression. For example, inoculation of *Arabidopsis* with *E. cloacae* UW4 increased the transcription of genes involved in cell division and proliferation while down-regulating genes involved in ethylene-induced stress and plant defense signaling (Hontzeas et al. 2004). These results make sense since bacteria would need to increase cell growth and proliferation in order to increase root length.

Bacterial Mechanisms of Colonization

The molecular mechanisms used by bacteria to colonize the interior of plants are still largely unknown. Observations that different bacteria colonize different plants to varying degrees indicate that colonization of the plant interior is an active process that is controlled by genetic determinants on both sides of the interaction (Dong et al. 2003b). Most studies focus on reporting the internalization of bacteria and their general location on and within plants, though some have linked bacterial colonization to specific genes. While many studies have been done on plant growth-promoting bacteria or plant pathogens, similar colonization mechanisms may be utilized by human pathogens.

Evidence for the host regulation of endophytic colonization by enteric bacteria came from the observation that the interior of an ethylene-insensitive mutant of *M. truncatula* was colonized in much higher numbers than the wild type plant (Iniguez et al. 2005). Subsequently, ethylene addition to plants greatly decreased endophytic colonization by *Salmonella* and *K. pneumoniae* 342, which could be reversed by the addition of an ethylene action inhibitor (Iniguez et al. 2005). As ethylene triggers induced systemic resistance (ISR) in plants, *Arabidopsis* lines that vary in

both the ISR and systemic acquired resistance (SAR) defense responses were tested for endophytic colonization. The results showed that ISR and SAR reduce plant invasion of *Salmonella* strains but only ISR affects interior colonization by Kp342 (Iniguez et al. 2005).

Salmonella mutants deficient in the production of flagella or the type III secretion system showed increased endophytic colonization (Iniguez et al. 2005). These systems appeared to elicit SAR but not ISR. As Kp342 lacks both flagella and type III secretion systems (Fouts et al. 2008), it is not surprising that this strain's ability to colonize the interior of plants is influenced only by ISR, not by SAR. These results suggest why *Klebsiella* strains are so commonly found in the interior of plants (Chelius and Triplett 2000; Chelius and Triplett 2001; Engelhard et al. 2000; Kuklinsky-Sobral et al. 2004; Reiter et al. 2003; Rosenblueth et al. 2004; Surette et al. 2003). One feature of the *Klebsiella* genus is that they lack flagella and, as in Kp342, may commonly lack some secretion systems as well. As a result, *Klebsiella* probably elicit a lower response from plant defenses than do most other bacteria.

Microarray analysis revealed that several pathogenicity genes, including type II secretion genes and a gene involved in capsular polysaccharide synthesis, were induced in *S. enterica* serovar Typhimurium during colonization in response to lettuce root exudates (Klerks et al. 2007). These genes are believed to aid in attachment to the plant (Klerks et al. 2007). Differential expression of genes related to carbon utilization, including *OtsA* (trehalose-6-phosphate synthase) and *UhpC* (hexose phosphate utilization protein), was also observed (Klerks et al. 2007). These results, in combination with chemotaxis experiments showing that the bacteria move toward root exudates, suggest that root exudates may serve as carbon sources for *Salmonella* and may condition *S. enterica* for root attachment by triggering chemotaxis and turning on genes that could play a role in adherence (Klerks et al. 2007).

In the nitrogen-fixing endophyte, *Azoarcus* sp. strain BH72, an endoglucanase gene is important for bacterial colonization of rice roots (Reinhold-Hurek et al. 2006). This endoglucanase, encoded by *eglA*, shows high homology to endoglucanases from the phytopathogenic bacteria, *Xanthamonas campestris* and *Ralstonia solanacearum* (Reinhold-Hurek et al. 2006). This gene is transcriptionally activated in the presence of rice roots, at lateral root emergence sites and lateral root tips (Reinhold-Hurek et al. 2006). A mutant in this gene is unable to spread into rice shoots and exhibits reduced intracellular colonization (Reinhold-Hurek et al. 2006).

Motility is also known to play a role in endophytic colonization. For example, root colonization and migration of the bacteria to other parts of the plant were decreased in nonmotile *Salmonella enterica* mutants compared to the wild type strain (Cooley et al. 2003). In addition, twitching motility is important for endophytic colonization by *Azoarcus* (Bohm et al. 2007). The structural component genes of type IV pili, *pilT* and *pilA*, are responsible for twitching motility, and mutants in these genes abolished the twitching phenotype (Bohm et al. 2007). The *pilA* mutant was strongly reduced in both rhizosphere and endophytic colonization. The *pilT* mutant was also greatly reduced in endophytic colonization compared to wild type but was only 50% reduced in rhizosphere colonization. Thus the twitching phenotype is pivotal in endophytic colonization but less important in root surface attachment (Bohm et al. 2007). The type IV pili locus, *pilAB*, is also important to plant colonization, with both genes required for attachment to rice roots by *Azoarcus* (Dorr et al. 1998). The *pilA* gene encodes a short pilin precursor whereas *pilB* shares similarity with FimF, a type I fimbria from *E. coli* (Dorr et al. 1998). The similarity of plant colonization genes to pathogenicity factors in human pathogens indicates these bacteria may be capable of colonization of their hosts by similar mechanisms.

Two partner secretion systems (TPS) also play roles in plant colonization by bacteria. In TPS, a TpsA family exoprotein with specific conserved secretion signals is transported across the membrane by a TpsB family channel-forming transporter that recognizes the secretion signal (Jacob-Dubuisson et al. 2001). The plant pathogen, *Erwinia chrysanthemi*, possesses a TPS, *hecAB*, in which *hecA* encodes an adhesin (Rojas et al. 2002). A mutant in this *hecA* gene had reduced attachment, cell aggregate formation, and virulence on *Nicotinia clevelandii* (Rojas et al. 2002). Homologs of this gene appear in both plant and animal pathogens and seem to be universal among necrotic plant pathogens (Rojas et al. 2002). In the mutualist *P. putida* KT2440, a TPS, *hlpBA*, is involved in seed colonization and iron uptake (Molina et al. 2006). The *hlpA* gene encodes a secreted protein similar to iron-regulated hemolysins and HlpB is responsible for its transport across the membrane (Molina et al. 2006). Mutations in this operon result in reduced colonization of corn seed and root attachment, demonstrating that the HlpA protein plays a role in plant attachment (Molina et al. 2006).

There are other factors involved in bacterial attachment and plant colonization. In *P. putida* KT440, a cell density-regulated gene, *ddcA*, is involved in corn seed colonization (Espinosa-Urgel and Ramos 2004). The function of this gene is still unknown, but *ddcA* mutants are reduced in seed adhesion whereas complementing the gene restores adhesion (Espinosa-Urgel and Ramos 2004). *ddcA* expression is induced both by cell density and by seed exudates, indicating that both bacterial and plant signals can regulate colonization (Espinosa-Urgel and Ramos 2004). Of interest in regard to human pathogens in plants, *ddcA* is a member of a conserved protein family found in many prokaryotes, including a gene expressed during macrophage invasion by *S. enterica* serovar Typhimurium and a phosphate starvation gene, *psiE*, in *E. coli* (Espinosa-Urgel and Ramos 2004).

An argument has been made for passive entry of *E. coli* into lettuce tissues based on the observation that FluoSpheres enter these plants at a similar rate and in similar numbers to *E. coli* cells (Solomon and Matthews 2005). However, comparison of living cells to inanimate objects can be difficult to interpret. Just because the rate and number of entry is the same between the two does not mean that the modes of entry are the same. In addition, plant defenses could be repressing the number of viable *E. coli* in plants, which can result in similar numbers of FluoSpheres and cells within the plant. FluoSpheres could thus be entering passively at a lower rate while *E. coli* cells are battling against plant defenses to maintain a relatively low number of viable cells within the plant. The simplest first experiments in this area would be to determine the effects of plant defense elicitors such as ethylene or salicylic acid on *E. coli* endophytic colonization of lettuce plants.

Surface adhesion plays a large role in pathogen colonization of plants. Several genes and mechanisms have been identified as being involved in attachment of human pathogens to plants. These mechanisms include curli, fimbriae, adhesins, and capsule production (Hassan and Frank 2004; Jeter and Matthysse 2005; Torres et al. 2005). When the curli biosynthesis regulatory gene, *mlrA*, or the curli biosynthesis operon, *csgA*-*G*, were introduced into *E. coli* K12, it conferred the ability to attach to alfalfa, while deletion of the *csgA* gene in diarrheagenic *E. coli* O157:H7 strains did not reduce attachment compared to wild type. Thus, the authors concluded that although curli may confer plant attachment to *E. coli* K12, they are not necessary for attachment by pathogenic *E. coli* as it may have multiple mechanisms involved in attachment (Jeter and Matthysse 2005). Transferring the adhesin genes *cah* and *aidA1* or the *csg* loci encoding fimbria formation into a nonbinding *E. coli* K12 also increased its ability to bind to alfalfa seeds and sprouts, though these genes did not have an additive effect on increasing

attachment when introduced together (Torres et al. 2005). As was observed with curli genes, deletion of these genes from a strain of *E. coli* O157:H7 did not alter binding to alfalfa; the only gene found to be required for attachment was *ompA* (Torres et al. 2005). These results suggest adhesins that mediate attachment in *E. coli* O157:H7 have overlapping and redundant functions (Torres et al. 2005). In addition, capsule production has also been associated with the ability of *E. coli* O157:H7 to attach to lettuce (Hassan and Frank 2004).

Common Virulence Factors in Plant and Animal Pathogens

Several studies have revealed the presence of similar virulence- and pathogenicity-related genes within both plant and animal pathogens. *Pseudomonas aeruginosa* PA14 has served as a model organism for this work as it is a broad host range pathogen. This opportunistic mammalian pathogen has been demonstrated to be a facultative pathogen of the model plant, *A. thaliana*, causing local and systemic infection resulting in plant death (Plotnikova et al. 2000). The bacteria were able to enter through stomata and wounds on leaves, colonize intercellular spaces, move basipetally along the vascular parenchyma, and disrupt plant cell walls and membrane structures (Plotnikova et al. 2000).

Several specific genes have been identified in *P. aeruginosa* as common bacterial virulence factors. In the process of developing and testing plant models to help elucidate virulence mechanisms of PA14, 9 mutants isolated for reduced virulence to plants were also found to be required for complete virulence on a burned mouse model, several of which correspond to genes with no known function (Rahme et al. 2000). Some of the genes identified in this study include *toxA*, *plcS*, *dbsA*, *hrpM*, *gacA*, and *gacS* (Rahme et al. 2000). Of these genes, *toxA*, *plcS*, and *gacA* were previously found to be common virulence factors necessary for causing disease in both *Arabidopsis* and a mouse model (Rahme et al. 1995). Two of these common virulence factors were the exported proteins, endotoxin A and phospholipase S.

Endotoxin A, encoded by *toxA*, inhibits protein synthesis (Iglewski and Kabat 1975; Ohman et al. 1980), and phospholipase S, encoded by *plcS*, attacks eukaryotic membranes (Ostroff and Vasil 1987).

The *dsbA* gene encodes a periplasmic disulfide bond-forming enzyme and may function to affect periplasmic virulence-related proteins (Bardwell et al. 1991; Rahm et al. 2000). DsbA has been observed to be important to the pathogenicity of both human and plant pathogens, including *Shigella flexneri*, *Vibrio cholera*, and *E. chrysanthemi* (Peek and Taylor 1992; Shevchik et al. 1995; Watarai et al. 1995). The *hrpM* gene is homologous to the *E. coli* gene, *mdoH*, involved in membrane-derived oligosaccharide synthesis (Loubens et al. 1993). Though its role in *E. coli* pathogenesis is not understood, deletions in *hrpM* in *P. syringae* pv. *syringae* reduces virulence, abolishing plant disease symptoms (Anderson and Mills 1985). *GacA* and *GacS* encode proteins from a two-component family regulator, in which GacS is the sensor kinase and GacA is its response regulator (Hrabak and Willis 1992; Laville et al. 1992). The *gacA* gene functions as a transcriptional regulator of pathogenicity genes encoding extracellular products (Rich et al. 1994). The observation that *gacA* functions at the regulatory level suggests that these common virulence factors share common regulatory mechanisms (Rahme et al. 1995). Mutants in these genes in *P. syringae* are decreased in lesion formation on bean (Hrabak and Willis 1992; Hrabak and Willis 1993; Rich et al. 1994). Overall, this study found there are several universal bacterial virulence mechanisms that are highly conserved in *P. aeruginosa* and are used to infect evolutionarily divergent hosts (Rahme et al. 2000).

Two modular pathogenicity islands have been found in *P. aeruginosa* PA14, with genes from a wide array of bacterial species and mobile elements (He et al. 2004). Most of the genes in these islands are homologous to ones found in human and plant pathogens, including type IV

group B pilus genes (He et al. 2004). Eleven of the genes in these islands are required for full virulence on both plants and animals (He et al. 2004). A large portion of the PAPI-1 island was similar to open reading frame (ORF) clusters in the phytopathogen *X. axonopodis* pv. *citri* and the human pathogen *S. enterica* serovar Typhi and may have been acquired by PA14 from these bacteria (He et al. 2004). Interestingly, 80% of the PAPI-1 island was unique, with ORFs that are not related to known proteins or functional domains, possibly representing a toolbox of pathogenesis gene variants for colonizing multiple hosts (He et al. 2004). Another gene present in *P. aeruginosa* involved in virulence on both plant and animals is *mucD*, which is orthologous to *degP*, a gene encoding a periplasmic protease and chaperone in *E. coli* with homologs to known virulence factors involved in stress responses by various species (Yorgey et al. 2001). A mutant in *mucD* had a reduced ability to cause disease in *Arabidopsis* and mice, and significantly lower growth in *Arabidopsis* (Yorgey et al. 2001). MucD is required for protection against environmental stresses, including temperature and oxidative stress, and also appears to be necessary for production of an extracellular virulence factor (Yorgey et al. 2001).

In addition to *P. aeruginosa*, several other bacterial pathogens colonize and cause disease in multiple hosts. *Burkholderia cepacia*, an important opportunistic pathogen in immunocompromised people (Coenye et al. 2001; Govan et al. 1996), can infect alfalfa (Bernier et al. 2003). In a study to develop a plant model for studying virulence of *Burkholderia*, most strains of *B. cepacia* that were virulent in alfalfa were also found to cause disease in a lung infection model (Bernier et al. 2003). Although not all *Burkholderia* strains were capable of virulence in alfalfa, those that were reduced in alfalfa virulence also caused less severe symptoms in the lung model (Bernier et al. 2003). This correlation reinforces the idea that some virulence factors may play a common role in both plant and animal infection. Another

opportunistic pathogen, *Enterococcus faecalis*, causes disease in *Arabidopsis*, attaching to the leaf surface, entering through stomata and wounds, colonizing the intercellular spaces, and causing rotting of cell walls and membrane structures (Jha et al. 2005). Two of five mammalian virulence factors tested, *fsrB*, a putative quorum sensing gene, and *sprE*, a serine protease, were important for plant pathogenesis, with mutants in these genes being strongly reduced in virulence and colonization on *Arabidopsis* (Jha et al. 2005). These observations suggest an "evolutionary crossover" of virulence factors between plants and animal pathogens (Jha et al. 2005).

In addition, the plant pathogen, *E. chrysanthemi*, adheres to human adenocarcinoma cells, and causes oxidative stress response followed by cell death (Duarte et al. 2000). A virulence factor important to *E. chrysanthemi* pathogenicity on human cells was a type III secretion system, since type III secretion mutants killed significantly fewer cells than wild type *Erwinia* (Duarte et al. 2000). In addition, *E. chrysanthemi* has a surface protein that shares immunological identity with the protein, intimin, which is required for virulence of pathogenic strains of *E. coli* (Duarte et al. 2000).

Since adherence to the surface of an organism is an important step in initiating infection, it is not surprising that some virulence genes in common between plant and animal pathogens are involved in attachment. For example, XadA, an outer membrane protein in *X. oryzae*, shares similarity to nonfimbrial adhesins from animal pathogens, including YadA from *Yersinia* and UspA1 from *Moraxella* (Ray et al. 2002). A region of similarity between YadA and XadA was repeated six times throughout the entire XadA protein (Ray et al. 2002). Also, homology modeling of XadA suggests it has a β-helix conformation, similar to the nonfimbrial adhesin pertactin, which is a virulence factor in the human pathogen *Bordetella pertussis* (Emsley et al. 1996; Ray et al. 2002). Infection of rice plants by *xadA* mutants was significantly reduced

compared to the wild type strain, demonstrating this gene does play a role in plant colonization (Ray et al. 2002). The *xadA* gene was induced on minimal media, indicating it is regulated by growth conditions (Ray et al. 2002). Regulation by growth conditions is not surprising for virulence genes, because they are needed only at the time of colonization.

Human pathogens also possess genes capable of mediating attachment to plants. In *Salmonella*, a screen for attachment mutants identified 20 genes necessary for *S. enterica* adherence to plants, 65% of which had no reported function (Barak et al. 2005). Some of these genes were also required for animal virulence (Barak et al. 2005). Specific genes identified in this screen include *agfB*, *agfD*, and *rpoS*, which are involved in production and regulation of curli and cellulose (Barak et al. 2005). In addition, *rpoS* regulates other virulence and adhesion factors, such as pili (Barak et al. 2005). Taken together, this information suggests that plants can serve as a secondary host and vector for *S. enterica* between animals (Barak et al. 2005).

Several other virulence genes in either plant or animal pathogens share similarity to each other. The *AvrA* gene in *S. enterica*, which encodes a target of the centrisome 63 type II secretion system, shares sequence similarity with AvrRxv, an avirulence determinant from the plant pathogen *X. campestris* pv. *vesicatoria* (Hardt and Galan 1997). AvrA is an effector protein translocated into host cells and it is suggested that AvrRxv and AvrA may be part of a family of effector proteins related to host/pathogen crosstalk (Hardt and Galan 1997). The presence of such common virulence genes between plant and human pathogens could explain how human pathogens are capable of colonizing plant hosts.

Genomes of Endophytic Bacteria

Genomes of endophytic bacteria are being sequenced, and analysis of this information is yielding insight into how these bacteria can inhabit the plant interior but not cause disease symptoms or cell death. The annotation of these genomes also provides targets for mutagenesis

to identify genes that may be involved in endophytic colonization. The first fully sequenced genome of a bacterial endophyte was that of the diazotrophic strain, *Azoarcus* sp. BH72. In BH72, there is a lack of toxin and virulence genes common in other plant pathogenic bacteria, including type III and IV secretions systems, presumably to avoid damaging the host tissue in the presence of high bacterial numbers within the plant (Krause et al. 2006). Also, these secretion systems can elicit plant defenses (Iniguez et al. 2005). Their absence may allow BH72 to avoid plant defenses. The BH72 genome has few phage or transposable elements, suggesting a low rate of lateral gene transfer, which is in contrast to the soil-associated *Azoarcus* sp. strain EbN1. This suggests that BH72 is adapted to a low stress, stable environment such as that found in plant tissue (Krause et al. 2006). BH72's lipopolysaccharide, exopolysaccharide, and capsule genes are more closely related to those from plant symbionts and pathogens than to its close relative, *Azoarcus* sp. EbN1 (Krause et al. 2006). In addition, the BH72 genome has a low degree of synteny with that of EbN1, while it possesses gene clusters that show significant synteny to plant pathogens, such as *P. aeruginosa* (Krause et al. 2006). Also in accordance with BH72's role as an endophyte, its genome appears to be adapted to low-nitrogen and low-iron environments, with high-affinity ammonia assimilation and transporter proteins, a full set of nitrogen fixation genes, and twice as many iron transporters as in EbN1, demonstrating that it can fix nitrogen and uptake both nitrogen and iron from the environment (Krause et al. 2006).

The genomes of other diazotrophic endophytes have either recently been completed or are nearly finished, including those of *K. pneumoniae* 342, *Gluconacetobacter diazotrophicus* PAl 5, and *Herbaspirillum seropedicae* Z67. Of these, the Kp342 genome has been published (Fouts et al. 2008). As expected, the full complement of *nif* genes was found in this genome. As in BH72, this genome has few mobile elements and many transport genes.

Given the plant origin of Kp342, it was surprising to find many animal virulence determinants in Kp342 that are present in clinical strains. Due to the presence of these genes, the animal virulence of Kp342 was tested in a mouse model and compared to a commonly studied clinical isolate. Kp342 did not colonize spleens or livers at a level similar to the clinical isolate. Kp342 did colonize lungs and kidneys where colonization is likely to be important in disease development. However, the level of Kp342 colonization of these two organs was 27 and 49 times lower in the kidneys and lungs, respectively, than with the clinical isolate. So although Kp342's pathogenicity in mice is less than that observed in a clinical isolate, the level of lung and kidney colonization is sufficiently high to suggest that Kp342 not be used in any agricultural application.

A preliminary report on the genome of the sugarcane endophyte, *H. seropedicae*, indicates that this endophyte also possesses nitrogen fixation genes, as well as 27 genes involved in iron transport (Pedrosa and Consortium 2005). In contrast to BH72, *H. seropediae* does have type III secretion genes, though both endophytes lack a type IV secretion system (Krause et al. 2006; Pedrosa and Consortium 2005). Important to plant microbe interactions, *H. seropedicae* possesses genes similar to *hrp* genes and *ysc* from *Yersinia pestis* (Pedrosa and Consortium 2005). In addition, *H. seropedicae* has genes similar to the type III secretion genes found in *Y. pestis* (Pedrosa and Consortium 2005). The presence of genes similar to those found in an animal pathogen indicates that these pathogens could also possess the ability to colonize plant tissue.

The *G. diazotrophicus* PAl 5 genome was recently deposited in GenBank. This bacterium was the first one definitively shown to provide fixed nitrogen to a grass species (Sevilla et al. 2001). As of this writing, a manuscript describing the genome has not been published. This

genome has two plasmids, 16.6 and 38.8 kilobases (kb) in size, and a 3.94 megabase (Mb) chromosome. Fully sequenced close relatives exist for all of the four endophytic genomes sequenced to date. These genomes should allow valuable comparative genomic analyses that should provide significant testable ideas on how plants interact with these bacteria.

Future Directions

Specific control mechanisms to reduce the number of human pathogens on or within raw produce have not been reviewed here. The objective here is to describe our understanding to date of the biology of endophytic colonization with a special emphasis on human bacterial pathogens. However, an understanding of endophytic colonization is useful not just in future control methods for human pathogens in edible plants but also to improve the growth and nutrition of plants through the use of diazotrophic endophytes. An understanding of the biology of endophytic colonization will result in a safer food supply and in more efficient food production. Control mechanisms are most valuable when they are coupled with a knowledge of the system. Fortunately, our knowledge of the biology of endophytic bacteria is increasing rapidly and should grow more given the availability of complete genomes. Perhaps the largest impediment to future progress in this area is our lack of understanding of plant defense mechanisms in food crops. Although the knowledge gained in recent years on plant defenses in *Arabidopsis* is very impressive, this knowledge needs to be translated to those plants that are consumed daily. Where that knowledge cannot be translated, we need to do the hard work to understand defense mechanisms in lettuce, tomato, spinach, and similar crops.

CHAPTER 3
CONFIRMATION OF THE SEQUENCE OF '*CANDIDATUS* LIBERIBACTER ASIATICUS' AND ASSESSMENT OF MICROBIAL DIVERSITY IN HUANGLONGBING-INFECTED CITRUS PHLOEM USING A METAGENOMIC APPROACH

Reproduced with permission from Tyler, H. L., Roesch, L. F. W., Gowda, S., Dawson, W. O., and Triplett, E. W. Confirmation of the sequence of '*Candidatus* Liberibacter asiaticus' and assessment of microbial diversity in Huanglongbing-infected citrus phloem using a metagenomic approach. Mol. Plant-Microbe Interact. (In press). In this study, H. L. Tyler performed all metagenomic analyses, most confirmatory polymerase chain reactions (PCRs), and wrote the manuscript. L. F. W. Roesch extracted the citrus deoxyribonucleic acid (DNA) and performed the automated method of ribosomal intergenic spacer analysis (ARISA), S. Gowda performed some confirmatory PCRs, and E. W. Triplett and W. O. Dawson provided input on the manuscript.

Introduction

Huanglongbing (HLB), also known as citrus greening, is a destructive and devastating disease of citrus causing great losses in citrus industries throughout the world (Bové 2006). This disease is characterized by yellow shoots, blotchy mottle on the leaves, and fruit that are small and lop-sided, with inverted coloring. HLB is transmitted between citrus trees by the psyllids, *Diaphorina citri* (in Asia and America) and *Trioza erytreae* (in Africa) (McClean and Oberholzer 1965; Capoor et al. 1967; Bové 2006). Progress in the study of this disease, including detection, treatment, and control, has been hindered due, in part, to the lack of effective culturing. Through microscopic examination and sequence analysis of polymerase chain reaction (PCR)-amplified 16S ribosomal DNA of infected tissue, it has been proposed that the causal agent of citrus greening is a phloem sieve tube-limited gram-negative bacterium that is a member of the alphaproteobacteria (Garnier et al. 1984; Jagoueix et al. 1994). More recent

studies have found HLB to be associated with three members of the genus '*Candidatus* Liberibacter': '*Ca.* Liberibacter asiaticus', '*Ca.* L. americanus', and '*Ca.* L. africanus' (Jagoueix et al. 1997; Teixeira et al. 2005).

In order to get around the obstacle of '*Ca.* Liberibacter spp.' being uncultured, attempts have been made to sequence the genome directly from infected tissue, in the hopes that knowledge of the genome will give better insight into how to culture this organism as well as how to combat it. Although '*Ca.* L. asiaticus' is at a low titer in citrus trees, it is present in much higher numbers when infected in periwinkle and Dodder (Ghosh et al. 1978; Garnier and Bové 1983). Therefore, DNA from infected plants in this system have been sequenced to obtain 8.56 and 14.7 kilobase (kb) segments of the '*Ca.* Liberibacter' genome using a genomic walking method (Doddapaneni et al. 2008; Lin et al. 2008). Recently, 34 contigs of '*Ca.* L. asiaticus' str. psy62 genomic sequence, ranging in size from 1.033 to 186.24 kb for a total of 1.2 megabases (Mb) of genomic sequence, has been deposited into GenBank (Accession number ABQW00000000). These genomic sequences were obtained from bacteria extracted from the psyllid vector (Duan et al. 2009). At the time of this writing, only 34 contigs were available for analysis, though the genome has been closed and is now published (Duan et al. 2009).

Therefore, in this article, we set out to confirm that the proposed '*Ca.* L. asiaticus' contigs are, indeed, associated with HLB symptomatic plants and not asymptomatic plants. Most HLB-infected trees in Florida are also infected with *Citrus tristeza virus* (CTV). We examined a CTV-free HLB isolate that was used to characterize the host range and symptoms of Florida HLB (Folimonova et al. in press). In addition, a non-PCR based metagenomic approach was used to investigate the microbial diversity within HLB-infected citrus tissue and whether there are any other pathogens present.

Results

Verification of '*Ca.* L. asiaticus' str. psy62 Contigs

In order to confirm the '*Ca.* L. asiaticus' contigs deposited in GenBank belong to the causative agent of HLB, it was necessary to verify that these contigs are present only in infected citrus tissue and not in healthy tissue. To that end, PCR reactions to assay for the presence of each contig were performed with genomic DNA from healthy and infected *Citrus sinensis* tissue. Before PCR, the 34 putative '*Ca*. L. asiaticus' contigs on GenBank were assembled and reduced to 32, with NZ_ABQW01000005 assembling with NZ_ABQW01000033 and NZ_ABQW01000022 assembling with NZ_ABQW01000031. Of these 32 contigs, 30 were found to have bands of the expected size range in infected tissue, but not healthy tissue (data not shown), confirming that these sequences are, indeed, associated with HLB symptoms and, therefore, from the likely causative agent of the disease. Of the 2 contigs that were not confirmed, there were either positive bands in both infected and healthy reactions or no positive bands in either reaction. These results could be explained by non-specific primer binding to regions in the *C. sinensis* genome.

Confirmation of '*Ca.* L. asiaticus' Contig Sequences

Because '*Ca.* L. asiaticus' contigs were detected only in infected tissue by PCR, a metagenomics approach was used to confirm that the entire sequence of each contig was present in HLB-infected phloem, further verifying that the sequences are from the causative agent of the disease. To do this, three next-generation sequencing platforms, 454, SOLiD, and Solexa, were used to generate a total of 13.6 gigabases (Gb) of metagenomic sequencing reads from HLB-infected citrus phloem (Table 3-1). With the large amount of sequence data obtained from HLB-infected phloem, high coverage of the '*Ca.* L. asiaticus' genome and, therefore, the '*Ca.* L. asiaticus' contigs, was expected. To determine the level of '*Ca*. L. asiaticus' coverage, all

metagenomic sequencing data was run in a reference assembly against the '*Ca*. L. asiaticus' contigs. Of all phloem sequences, 0.23% matched the '*Ca*. L. asiaticus' contigs, with coverage ranging from 20 to 60-fold per contig (Table 3-2). In addition, sequences from HLB-infected phloem spanned the entire length of each contig, with 99.69 to 100 % of each contig matching the consensus sequence generated from the reference assembly (Table 3-3).

The dataset from each high-throughput sequencing method was also run against the '*Ca.* L. asiaticus' contigs separately in order to determine if there were any biases due to sequencing method. One of the major differences between the results of the three sequencing methods appears to be due to number of reads, with the 58.8 Mb of 454 sequences only yielding 0 to 0.341-fold coverage of the '*Ca*. L. asiaticus' contigs (Table 3-2). In contrast, reference assemblies with Solexa and SOLiD data, which provide significantly more sequences (1.78 and 11.8 Gb), yielded 10.78- to 28.44-fold and 5.62- to 31.29-fold coverage, respectively, of the '*Ca*. L. asiaticus' contigs (Table 3-2). When looking at the consensus sequences generated by these two reference assemblies, Solexa sequences covered 98.9 to 100 % of each '*Ca.* L. asiaticus' contig and SOLiD data covered 82.39 to 100 % (Table 3-3). Of the three sequencing methods, Solexa yielded slightly better fold coverage of the '*Ca*. L. asiaticus' contigs overall. For example, 23 of the 32 Liberibacter contigs had higher fold coverage when assembled against Solexa data than with 454 or SOLiD (Table 3-2). Specifically, contig NZ_ABQW01000028 had 5.62-fold coverage with SOLiD sequences, but when assembled with Solexa sequences, fold coverage of this contig almost tripled, going up to 15.03 (Table 3-2). There were also cases where SOLiD data yielded higher fold coverage of a contig but the increase was not as much. In one such case, contig NZ_ABQW01000026 had 18.077-fold coverage with SOLiD data and 10.862 with Solexa (Table 3-2). Sequences from 454 did not approach the level of coverage

seen using SOLiD or Solexa for any of the contigs, mostly due to the reduced number of sequences from 454 compared to the other two. For example, 454 yielded 0.25 million reads while SOLiD and Solexa yielded 337 and 49 million reads respectively. However, overall, sequences from each of the 3 methods complemented gaps missed by the other two because, when analyzed separately, only 4 of the 32 '*Ca*. L. asiaticus' contigs had consensus sequences that covered 100% of their length but, when all data were assembled together against the '*Ca*. L. asiaticus' contigs together, 11 of those contigs had a 100% match (Table 3-3). With such high fold coverage and consensus sequences covering the entirety of each contig, it was concluded that these contigs are present in the phloem.

Comparison to Alphaproteobacterial Relatives

In order to discount the possibility that the matches seen when comparing citrus phloem sequences to '*Ca.* L. asiaticus' contigs are due to the presence of close alphaproteobacterial relatives, all infected phloem sequence data was compared to 5 fully sequenced alphaproteobacteria deemed to be closely related to '*Ca*. L. asiaticus' after a BLAST of its contigs against an alphaproteobacterial database. The results of the reference assemblies with each of the 5 alphaproteobacteria showed that over 100-fold fewer sequencing reads match these alphaproteobacteria compared with '*Ca*. L. asiaticus' contigs (Table 3-4). For example, the closest match, *Rhizobium leguminosarum* bv. *viciae* 3841, had 8,012 reads match its genome while 900,124 reads matched the '*Ca*. L. asiaticus' contigs. In addition, sequencing reads from infected phloem yielded an average 26.34-fold coverage of the '*Ca.* L. asiaticus' contigs but none of the alphaproteobacteria had greater than 0.175-fold coverage (Table 3-4). Therefore, the bacterial DNA in phloem is specific to '*Ca.* L. asiaticus' and even close relatives do not match this DNA. This confirmation of '*Ca.* L. asiaticus' in symptomatic plant tissue further supports its proposed role as the causative agent of HLB.

Assessment of Bacterial Diversity

Automated ribosomal intergenic spacer analysis (ARISA) was performed for the initial assessment of bacterial diversity within HLB-infected phloem. Amplification with ARISA primers yielded 2 dominant peaks in healthy (740 and 1,330 basepairs (bp)) and 3 dominant peaks in infected (740, 1,330, and 1,600 bp) phloem (Figure 3-1). To determine the identity of these peaks, the ARISA PCR was run on a gel, and the bands extracted, cloned, and sequenced in a 96-well plate. The 740 and 1,330 bp bands were from *C. sinensis* chloroplast and mitochondria, respectively, while the 1,600 bp band unique to infected tissue was from '*Ca.* Liberibacter' 16S ribosomal ribonucleic acid (rRNA). Of all the sequences obtained from the 1,600 bp band, there was a 100-fold excess of '*Ca.* Liberibacter' sequences, indicating it as the most abundant bacterium present within the phloem.

Once the '*Ca.* L. asiaticus' contigs were confirmed in infected citrus tissue, metagenomic sequencing data from the infected phloem was analyzed to further assess whether '*Ca.* L. asiaticus' was the only bacterium present in this sample. Bacterial diversity in the phloem of *C. sinensis* plants infected with '*Ca.* L. asiaticus' was analyzed by searching total DNA sequences obtained from infected phloem for 16S rRNA gene sequences. Because plant chloroplasts are related to Cyanobacteria, plant chloroplast sequences were filtered out of the phloem sequences. To do this, the entire dataset was assembled against the *C. sinensis* chloroplast genome and nonassembled reads were taken for further analysis. Approximately 5.7 million of the total sequencing reads matched the chloroplast genome and were removed. The remaining 381,056,442 sequences were run in a reference assembly against the Ribosomal Database Project (RDP) database. Sequences from the 454, Solexa, and SOLiD data sets were run both separately and together. Each reference assembly yielded an output with the number of sequencing reads matching each reference in the database, average coverage of each 16S rRNA match, and length

of the consensus sequence matching each 16S rRNA gene. In each assembly, the 16S rRNA sequence from the database with the greatest number of matches to infected citrus phloem sequence was '*Ca.* Liberibacter'. In addition, the '*Ca.* Liberibacter' 16S rRNA gene had highest fold coverage and longest consensus sequence over the entirety of the gene compared with any other bacteria in the database (Table 3-5). In total, the phloem metagenome from all three sequencing platforms combined yielded a fold coverage of 23.17 of the '*Ca.* Liberibacter' 16S rRNA gene, with a consensus sequence of 1,212 bp in length (Table 3-6). The length of the '*Ca.* Liberibacter' 16S rRNA sequence in the database was 1,448 bp; therefore, 83.7% of the 16S rRNA gene was covered (Table 3-6). In contrast, the next highest match, the 16S rRNA gene from *Nitrospina gracilis*, had a fold coverage of 1.45 and a consensus sequence that only covered 7.15% of the gene. All other matches to members of the RDP database covered less than 14% of the total 16S rRNA gene. Even coverage over the entire length of the 16S rRNA gene is expected for any bacterium present in the phloem; therefore, '*Ca.* Liberibacter' is likely the only bacterium present. In addition, the short regions within the 16S gene of the other RDP matches had 90 to 100% identity with sequences in '*Ca.* L. asiaticus' and citrus chloroplast and mitochondria when compared against these references on GenBank. Therefore, it was concluded these matches were false positives. Given that it would be impossible eliminate all false-positive matches due to alignment and sequencing errors, a threshold fold coverage detection limit was determined. In order to calculate this detection limit, the mean and standard deviation were calculated from the fold coverage of every RDP 16S rRNA match. Any RDP match with a fold coverage greater than the mean plus 3 standard deviations (1.89) was deemed to be significant and above the threshold fold coverage that would be seen with any 16S rRNA gene assembling to remaining citrus and '*Ca.* Liberibacter' sequences due to misalignments or sequencing errors.

Because '*Ca.* Liberibacter' was the only RDP database member to have coverage above 1.89, it was deemed to be the only bacterium likely present in citrus phloem.

Similar results were seen when SOLiD and Solexa data were run against the RDP database separately, yielding a fold coverage of the '*Ca.* Liberibacter' 16S rRNA of 17.37 and 5.23, respectively (Table 3-6). When 454 sequences were run separately, fold coverage was 0.33, with a consensus sequence matching only 26.17% of the '*Ca.* Liberibacter' 16S rRNA gene (Table 3-6). This number is not entirely surprising because very few 454 reads matched any of the 16S rRNA sequences. Given that the average read length of the 454 sequences was 248 bp and only two of those reads matched the 1,448 bp '*Ca.* Liberibacter' 16S rRNA gene, a match to only 26% of the 16S rRNA is unsurprising. There was only one other species in the entire database that had a matching member when run against the 454 data and it had 100-fold lower coverage than '*Ca.* L. asiaticus' (Table 3-5). As a result, the 454 data had little effect on the results of citrus phloem community analysis using the RDP database. It is likely this is due to the low amount of sequence generated by 454 compared to the other two methods. It is clear, however, that both SOLiD and Solexa contributed to the community analysis, since the overall fold coverage determined by the reference assembly using all three sets of sequencing data was higher than the fold coverage found when analyzing these datasets separately (Table 3-6).

Other Reported HLB Associated Bacteria Not Present

In addition to '*Ca.* Liberibacter spp.', other bacteria have been reported to be associated with HLB infection in citrus, including *Propionibacterium acnes*, '*Ca.* Phytoplasma asteri', and pigeon pea witches'-broom phytoplasma (Davis et al. 2008; Chen et al. 2009; Teixeira et al. 2008). These reports are in contradiction to the results of the 16S rRNA analysis of the citrus metagenome, which yielded a fold coverage of 0.023 and 0.026 for '*Ca.* P. asteri' and *P. acnes*, respectively, which corresponds to one 35 bp sequencing read matching the 16S rRNA gene. To

further confirm the presence or absence of these bacteria in the current study, reference assemblies of the metagenomic sequences were run against the genomes of these bacteria available on GenBank. In both cases, the fold coverage of these genomes was more than 100-fold lower than the coverage of '*Ca.* L. asiaticus' (Table 3-7). These results indicate that '*Ca.* P. asteri' and *P. acnes* are not present in the sample; otherwise, much higher coverage of these genomes would be expected if these bacteria were present in citrus phloem. In comparison, even the non-HLB-associated actinobacterial relative of *P. acnes*, *Micrococcus luteus*, had 0.038-fold coverage when assembled with the citrus metagenomic sequences, double the fold coverage of *P. acnes* (Table 3-7). Seeing a non-HLB associated bacterium with a higher fold coverage than a bacterium reportedly associated with HLB indicates that the genome of any bacteria will match such a large amount of metagenomic sequences due to alignment with conserved genes. Upon closer inspection, 84 to 99% of the sequencing reads that matched the genomes of these three bacteria aligned to regions in the 16S and 23S rRNA genes. If *P. acnes* was truly present in the sample, a more even distribution of reads aligning over the genome would be expected. It is likely that a large number of these sequences matching the 16S and 23S rRNA genes are derived from the rRNA genes of citrus chloroplast and mitochondria, which would be present in citrus tissue in higher numbers and, therefore, inflate the fold coverage of any bacterial genome compared to the metagenome. To test this idea, rRNA regions in the reference assemblies with many metagenomic sequence matches were found to have greater than 90% and, in some cases, 100%, identity with sequences from the *C. sinensis* chloroplast genome. Furthermore, when the reference assemblies were repeated with reads that had the *C. sinensis* chloroplast genome filtered out, the number of matching bases and fold coverage of '*Ca.* P. asteri', *P. acnes*, and *M. luteus*, as well as the genomes from 5 reference alphaproteobacteria,

were significantly reduced (Table 3-8). In contrast, the fold coverage of '*Ca.* L. asiaticus' contigs remained above 26 even with chloroplast sequences removed. These analyses validate the observation that the majority of sequences matching these genomes were chloroplast in nature and these other bacterial species were not present in the phloem sample because any bacterial genome is likely to have a low level of fold coverage when compared against such a large number of metagenomic sequences due to the conserved regions in rRNA genes.

In order to test the hypothesis that '*Ca.* P. asteri' and *P. acnes* were not present in citrus phloem, a detection limit was calculated to determine the threshold level of sequencing matches required to be 99% confident that a bacterial genome is truly present in the sample. The detection limit of these metagenomic sequences was set at three standard deviations above the mean fold coverage of seven bacterial genomes referenced against the entire metagenomic dataset. At this level, any genome with a fold coverage of 0.29 or greater was considered to be above the background level of false positive sequence matches and deemed to be present in the citrus phloem. This corresponds to 0.0021% of all bases and reads in the metagenomic dataset, or approximately 8,144 of the 386,745,534 total reads and 285,030 bases out of the total 13.6 Gb of metagenomic data. '*Ca.* L. asiaticus' was the only genome analyzed with a fold coverage greater than this threshold. Neither '*Ca.* P. asteri' nor *P. acnes* had enough sequencing reads match their genomes to achieve a fold coverage approaching this limit. Therefore, '*Ca.* L. asiaticus' was the only one present within these phloem samples.

Approximately 0.23% of the 13.65 billion bases sequenced in this work are from '*Ca.* L. asiaticus'. Assuming a diploid genome size of citrus at 900 Mb, these data represent a 15-fold coverage of the diploid citrus genome or can be considered the full chromosomal content of 15 citrus phloem cells. Assuming a '*Ca.* Liberibacter' genome size of 1.21 Mb, the 26-fold

coverage of the '*Ca.* Liberibacter' genome obtained here can be considered to be the full chromosomal content of 26 '*Ca*. L. asiaticus' cells. This suggests that the phloem sample sequenced here possessed 1.7 '*Ca*. L. asiaticus' cells per phloem cell. Based on our detection limit of 0.29-fold coverage, the metagenomic approach described here could detect one '*Ca*. L. asiaticus' cell for every 52 phloem cells.

DNA Viruses and Viroids Not Present

Because Koch's postulates have not been completed for HLB, it is possible that other agents, such as a virus or viroid, may be involved in the development of disease symptoms. To determine whether any such agent is present in HLB-infected phloem, all metagenomic sequence data was compared against a virus and viroid database. When a reference assembly of greening infected phloem sequences against the viroid database was performed, no sequencing reads assembled, indicating that there are no known viroids present within the phloem. This result is not unexpected as the phloem sample sequenced was of DNA extracted directly from the plant tissue and no reverse-transcriptase steps to obtain RNA sequences were performed. The reference assembly against the virus database did yield matches, but none showed good coverage. The highest fold coverage to any member of the virus database was 2.04-fold but the consensus sequence only spanned less than 1% of the reference that it matched. Considering the short nature of the consensus sequence and the fact that the match was to Hepatitis, a human virus, it is unlikely that this virus is present in the phloem sample. While the amount of sequencing in this study may not be sufficient to detect low abundance viral particles, the amount of sequence obtained for this work should be sufficient to detect virus particles that are sufficiently abundant to cause disease. Therefore, DNA viruses are not likely to play a role in HLB in the phloem.

High Coverage of Citrus

The DNA sequenced for this metagenomic analysis was extracted from total citrus phloem; therefore, it was expected that the dataset would include a high coverage of the *C. sinensis* genome in addition to the genomes of any microorganisms present. In order to confirm this, all phloem sequences were run through a reference assembly against a citrus expressed sequence tag (EST) database. Of the 34,505 ESTs in the database, 79.85% had hits from the phloem sequence, with an average fold coverage of 13.19 (results not shown).

Discussion

Most of the research results that associate '*Ca.* Liberibacter spp.' to HLB diseases in citrus involve microscopic examinations, cloning and sequencing of PCR amplified DNA from infected tissue, and other PCR-based techniques. In this study, a PCR independent metagenomic approach was taken to confirm the presence of '*Ca*. L. asiaticus' contigs, determine if other organisms may be present, and assess the microbial diversity in HLB-infected phloem. Prior to confirming the role of '*Ca*. L. asiaticus' in infected citrus phloem, it was first necessary to verify that the contigs for this strain's genome in GenBank were only present in infected tissue. To that end, PCR performed with primers designed to each of the '*Ca*. L. asiaticus' contigs confirmed that 30 of the 32 contigs were present in HLB-infected but not healthy citrus tissue.

Once it was confirmed that the '*Ca*. L. asiaticus' contigs were present only in HLB-infected tissue, the presence of those contigs in infected citrus phloem was examined by metagenomics. In this analysis, all contigs were present in infected phloem with fold coverage ranging from 20.642 to 60.067. Such high fold coverage of these contigs from HLB-infected phloem further confirms the validity of these contigs belonging to the genome of '*Ca*. L. asiaticus' and a causative agent of HLB. It is unlikely that the high coverage of '*Ca*. L. asiaticus' is due to the presence of another alphaproteobacterium in the tissue since a selection of

genomes from 5 different alphaproteobacterial relatives to '*Ca.* Liberibacter spp.' did not have greater than 0.2-fold coverage.

Because Koch's postulates have not been completed for '*Ca.* Liberibacter spp.' in HLB disease, it is possible that another pathogen or pathogens may play a role in disease development. For this reason, the composition of the microbial community in infected citrus phloem was also investigated by assembling the citrus phloem metagenome sequences against the 16S rRNA database from RDP. This analysis was run with data from three different high-throughput sequencing methods separately as well as with all the data combined and, in each case, the '*Ca*. L. asiaticus' was found to be the only bacterium with significant coverage of its 16S rRNA gene. For all other bacteria in the RDP database, metagenomic sequences matched no greater than 13.5% of the 16S gene or a fold coverage of 1.45. If any other bacteria were present within the citrus phloem, a more even coverage over the entire 16S rRNA gene would be expected because the total DNA extract from infected phloem was sequenced without using any PCR amplification that would create biases. Therefore, it can be said with confidence that the matches of the contigs and 16S rRNA gene with the metagenomic dataset from HLB-infected phloem are indeed due to the presence of '*Ca*. L. asiaticus' in infected phloem and not another bacterium. In addition, no known DNA virus was identified in the metagenomic analysis, although, in the field, most HLB-infected trees are also infected with CTV. However, the samples examined were chosen because they are free of CTV. We have examined the interaction of CTV and HLB and found little or no synergistic interactions (Folimonova et al. in press). Therefore, viruses were discounted as playing a role in HLB.

The results from this study differ slightly from others assessing the microbial diversity of HLB-infected citrus plants. In the current study, the only bacterium found to be present in

infected citrus was '*Ca*. L. asiaticus' whereas Sagaram and associates (2009) found a large diversity of microorganisms, comprising 47 orders of bacteria. However, two very different samples were analyzed in these studies. Sagaram and associates (2009) were looking at entire leaf midribs whereas the current study looked at phloem scraped from the inside of *C. sinensis* bark that was peeled from growing trees. The entire midrib includes the apoplast, which is known to harbor bacterial endophytes. It is likely that a larger array of bacteria are present in leaf tissue than are actually inside the phloem cells of the plant vascular system. The causative agent of HLB is thought to be restricted to phloem sieve tubes; therefore, this tissue is ideal for examining HLB-associated DNA sequences. Because '*Ca.* Liberibacter spp.' is intracellular within phloem cells, being spread systemically from the site of infection to other tissues, there will be fewer bacteria present in this environment as any present would have to bypass more plant defense mechanisms to enter the cell. Although midribs do contain phloem tissue, other leaf tissue is also present in such preparations and would contain additional bacteria that would inflate the estimate of microbial diversity directly in contact with '*Ca*. L. asiaticus'.

In addition, Chen and associates (2009) reported a strain of '*Ca.* P. asteri' associated with HLB-infected citrus that was also not identified in the current study. Because that study also examined leaf midribs as opposed to phloem cells from the bark, differences in the bacteria identified are not surprising. Although this study also reported observing phytoplasma in phloem sieve tube cells from midribs using electron microscopy, the majority of the microscopic examinations were performed in periwinkle, which showed an enrichment of the bacterium compared to citrus. Furthermore, the identification of phytoplasmas in sieve tubes was done on the basis of cell wall thickness, without any molecular means of confirming their identity such as a phytoplasma-specific antibody. Thus, it is not certain that phytoplasmas are present in the

sieve tube cells of citrus. However, these results do not preclude the possibility that phytoplasmas can cause citrus greening-like symptoms in citrus in some areas. In this work, citrus is shown to have these symptoms with '*Ca.* Liberibacter spp.' exclusively.

In another study, Davis and associates (2008) looked specifically at phloem tissue in an attempt to culture '*Ca*. L. asiaticus' and obtained a co-culture of '*Ca.* L. asiaticus' with actinobacteria related to *P. acnes* based on 16S rRNA analysis, with a 1,006 bp 16S rRNA sequence from the co-culture having 100% identity to *P. acnes*. These observations are in contrast to the results of the present study because there was no significant coverage of the *P. acnes* genome in the metagenomic sequences. Although '*Ca*. L. asiaticus' was reported to have been co-cultured with *P. acnes*, it took 31 culturing attempts to obtain a second co-culture with these bacteria and 12 additional attempts to obtain a third co-culture from HLB-infected plants (Davis et al. 2008). The number of repeated attempts necessary to obtain co-cultures with *P. acnes* from HLB-infected tissue indicates this bacterium may not be a common inhabitant of citrus. If *P. acnes* was commonly found in citrus plants and capable of supporting '*Ca*. L. asiaticus' in co-culture, a higher abundance of *P. acnes* DNA would be expected in these plants. In addition, Davis and associates (2008) used *P. acnes* specific primers to detect the bacterium in DNA extracts from citrus but only 3 healthy plants were reported to be positive for *P. acnes*. Also, whereas Davis and associates (2008) state they identified *P. acnes* in their co-cultures, they did not specify whether PCR was performed to identify *P. acnes* in DNA directly isolated from HLB-infected plants. Furthermore, *P. acnes* is a commensal of human skin (Brüggemann et al. 2004), and its identification in plant tissues may be the result of contamination while handling the samples. In this work, *P. acnes* was not found in the metagenomic sequences and may not be associated with '*Ca*. L. asiaticus' in planta.

Given the large amount of sequencing data generated by the three high-throughput sequencing methods, there is a high likelihood that the genome of any bacteria present in the phloem sample would have been detected. To test this, a detection limit of 3 standard deviations above the mean observed fold coverage of a bacterial genome was used to distinguish between the false and genuine sequence matches to the '*Ca*. L. asiaticus' genome. Only the '*Ca*. L. asiaticus' genome sequences met this standard. In addition, the level of 16S rRNA coverage by other bacteria was also much lower than expected for a bacterium to be present in these samples. In both cases, when looking at entire genomes or specifically at 16S rRNA genes, '*Ca.* Liberibacter' was the only bacterium identified above these detection limits. Furthermore, fold coverage of the '*Ca.* Liberibacter' 16S rRNA and genome were 12 and 90 times higher than these limits, respectively. Because '*Ca*. L. asiaticus' occurs in very low concentrations in infected citrus (Tatineni et al. 2008) and was identified at such high levels above the detection limit, we are confident that any other low-abundance bacteria present in the tissue would have been detected by this method.

This article presents a new approach to evaluating the microbial diversity present in citrus tissue by a PCR-independent metagenomic method. This method has the advantage of avoiding biases based on PCR amplification and cloning (Polz and Cavanaugh 1998). In addition, by using high-throughput sequencing methods, a greater sequencing depth of the entire citrus phloem sample was obtained at a lower per-base cost than traditional Sanger sequencing. Similar metagenomic studies have also been applied to investigate the causative agents of other diseases. Using 454 pyrosequencing, colony collapse disorder (CCD) in honey bees (*Apic mellifera*) was found to be associated with Israeli acute paralysis virus (IAPV), an unclassified dicistrovirus (Cox-Foster et al. 2007). In that study, results from the metagenomic analysis

contrasted with previously proposed causative agents of CCD. Although IAPV sequences were found only in CCD cases, *Nosema ceranae* and an iflavirus sp. proposed to be associated with the disease were identified in both CDD and non-CCD cases (Cox-Foster et al. 2007). As a result, this demonstrates how the metagenomics approach can be used to pinpoint disease agents free from culturing, PCR, and cloning biases.

In summary, this study used a culture- and PCR-independent metagenomic approach to verify the validity in citrus of the reported '*Ca*. L. asiaticus' contigs that were obtained from infected psyllids, determine that it is the major and perhaps the only bacterium present in the phloem of infected citrus phloem, and thereby substantiate its proposed role as the likely causative agent of HLB. With the decreasing costs in sequencing due to high-throughput methods and the level of sequence coverage of samples analyzed, metagenomics will become a most effective method for studying uncultured plant pathogens in the future. This will allow researchers to examine a broader view of the complexity of organisms present in the diseased state.

Materials and Methods

Sample Preparation and DNA Extraction

Eleven citrus plants, including *C. volkameriana*, *C. aurantifolia* cv. Swingle, *C. limon* cv. Eureka, 'Hirado Buntan Pink,' and *C. sinensis* cv. Valencia, symptomatic for citrus greening, were sampled. Branches were cut from plants and surface sterilized by wiping with tissue paper and 70% ethanol. The bark, which was 'slipping' in growing plants, was removed and the phloem cells were scraped off by using a sterile, disposable scalpel. DNA was extracted directly from the phloem cells.

DNA was isolated from the collected phloem tissue samples (approximately 150 mg) using the FastDNA Kit (Qbiogene, Inc., Calif.). After the DNA extraction, samples were purified with

the DNeasy Tissue kit (Qiagen, Valencia, CA.) following the manufacturer's instructions. This procedure was repeated to obtained enough DNA for 454, SOLiD, and Solexa sequencing.

PCR to Confirm Validity of '*Ca*. L. asiaticus' Contigs in GenBank

DNA extracted from healthy and infected *C. sinensis* cv. Valencia (as described above) was used as template in PCR reactions with primer pairs for each '*Ca.* L. asiaticus' contig in GenBank. Primers were designed to amplify a 0.5, 1, or 2 kilobase (kb) fragment from each contig using Primer3 (v. 0.4.0) (http://frodo.wi.mit.edu/) (Table 3-9). PCR reactions contained 4 µL of 5X HF Phusion Buffer (New England Biolabs), 200 µM of each dNTP, 2.5 µM of each primer, 0.4 U of Phusion polymerase, and approximately 50 ng DNA template in a final volume of 20 µl. PCR reactions were held at 98˚C 30 seconds (s); followed by 30 cycles of 98˚C for 10 s, 60˚C for 30 s, and 72˚C for 1 minutes (min); with a final extension at 72˚C for 10 min. PCR products from healthy and HLB-infected samples were resolved on a 1% agarose gel and examined for bands in the expected size range.

ARISA Analysis, Cloning and Sequencing for Initial Assessment of Bacterial Diversity and to Confirm Presence of '*Ca*. L. asiaticus'

Bacterial community composition was assessed by ARISA (Fisher and Triplett 1999). ARISA is a molecular technique for constructing bacterial community fingerprints based on the length heterogeneity of the intergenic transcribed spacer region of bacterial rRNA operons (Bosshard et al. 2000). Treating the elements of ARISA profiles as operational taxonomic units allows for bacterial community comparisons. In this study, ARISA profiles were assumed to be indicative of bacterial community composition, and differences in ARISA profiles were assumed to reflect variation in the composition of the respective bacterial communities.

PCR reaction mixtures contained 5 µl of 10X PCR buffer (Promega), 200 µM of each dNTP, 100 µM of each primer, 2.5 U of *Taq* polymerase, and approximately 100 ng of DNA

template (estimated by using NanoDrop ND-1000 spectrophotometer) in a final volume of 50 µl. The primers used were 787F (5'-ATTAGATACCCNGGTAG-3') (Roesch et al. 2007) and L-D-Bact-132-a-A-18 (5′-CCGGGTTTCCCCATTCGG-3′) (Ranjard et al. 2001). Reaction mixtures were held at 94°C for 3 minutes; followed by 30 cycles of amplification at 94°C for 45 s, 55°C for 1 min., and 72°C for 2 min; and a final extension of 72°C for 7 min.

Sample fragments were then discriminated by using on-chip gel electrophoresis with the Agilent 2100 Bioanalyzer and DNA LabChip Kit 7500. Briefly, a ladder with known fragment sizes is loaded and a standard curve of migration time versus fragments size is plotted. The size of each fragment in the sample is then calculated by taking into account the migration times measured. Lower and upper marker standards are also run with each sample in order to align the ladder data with data from the sample wells. Data is translated into gel-like images (bands) and electrophelograms (peaks).

After the dominant peaks were detected, the PCR products were loaded in a 1.2% agarose gel and the bands were identified and excised. The bands were purified using QIAEX II Gel Extraction Kit (Qiagen, Valencia, CA) following the manufacture's instructions. The purified fragments were cloned into a TOPO TA Cloning (Invitrogen, Carlsbad, CA). Plasmids were purified with the QIAprep spin miniprep kit (Qiagen, Valencia, CA) and gene fragments were cycle sequenced using T7 primer in a DYEnamic ET terminator cycle sequencing kit (GE Healthcare), on a PTC200 thermocycler (BioRad) and run on a 96 well MegaBACE 1000 capillary sequencer (GE Healthcare). Once the presence of '*Ca.* Liberibacter spp.' was confirmed by sequencing, DNA from 1 positive tree was used for further analysis. Sequences of ARISA bands were deposited onto GenBank under accession numbers GQ254546 to GQ254633.

ARISA Sequence Analysis and Classification

The sequences obtained for each excised band were initially analyzed by using VecScreen, an on-line tool for identifying segments of a nucleic acid sequence that may be of vector origin (http://www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html). After manual edition for the elimination of potentially chimeric sequences, phylogenetic analyzes were conducted using MEGA version 4 (Tamura et al. 2007). The evolutionary distance among the sequences was inferred using the Neighbor-Joining method (Saitou and Nei 1987), and a bootstrap test (500 replicates) was conducted in order to calculate the confidence limits of the phylogeny (Felsenstein 1985). All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option).

High-Throughput Sequencing

Three high-throughput sequencing methods were used to assess the bacterial diversity of a symptomatic *C. sinensis* cv. Valencia plant. DNA from one '*Ca.* Liberibacter sp.' positive plant was used for 454 pyrosequencing, Solexa, and SOLiD sequencing. Analyses were performed with DNA extracted from an infected citrus tree positive for citrus greening according to the ARISA analysis, cloning, and sequencing of the DNA fragments. For 454, 10 µg of DNA were sent to the Interdisciplinary Center for Biotechnology Research (ICBR) at the University of Florida (http://www.biotech.ufl.edu/), prepared according to the manufacturers instructions and sequenced using the 454 Genome Sequencer FLX System (454 Life Sciences, Branford, CT). An additional 10 µg of DNA were sent to ICBR at the University of Florida and used for SOLiD sequencing according to the manufacturer protocols (Applied Biosystems, Carlsbad, CA). For Solexa sequencing, 10 µg of DNA were sent to the Center for Genome Research and Biocomputing at Oregon State University (http://www.cgrb.oregonstate.edu/) where it was

prepared using the manufacturer's standard protocols and run on the Illumina 1G Genome Analyzer (Illumina, Hayward, CA). All metagenomic sequences are available from the National Center for Biotechnology Information (NCBI) Short ReadArchive, accession numbers SRR016809, SRR016816, and SRR017902.

Sequence Analysis

Reads obtained from each sequencing method were analyzed using CLC Genomics Workbench Version 3.2 (www.clcbio.com). Parameters used were 99% similarity for 454 reads, a limit of 1 for Solexa and SOLiD, and a mismatch cost of 3. All other parameters were set to the program's default. Analyses of the data from each sequencing method were run individually and pooled together. To determine microbial community composition within *C. sinensis* phloem, all sequence data was first run through a reference assembly against the *C. sinensis* chloroplast genome from NCBI (NC_008334) to eliminate false positives from chloroplast ribosomal RNA. Sequencing reads that did not assemble to the chloroplast genome were used in a reference assembly against the 16S database from the Ribosomal Database Project (http://rdp.cme.msu.edu). The reference assemblies were done on the entire dataset as well as individually for 454, Solexa, and SOLiD data. To determine if any viruses or viroids were present in the phloem of the greening infected plant, all sequencing reads were run in reference assemblies against viroid and virus databases. The viroid database was composed of all 39 completed viroid genomes from NCBI and the virus database was composed of all annotated virus genomes from the Viral Bioinformatics Resource Center (http://athena.bioc.uvic.ca/).

To estimate the level of '*Ca.* L. asiaticus' genome coverage present in the phloem sequences, a reference assembly of the data was run against all '*Ca.* L. asiaticus' str. psy62 contigs available in GenBank (NZ_ABQW00000000). Four reference assemblies were performed with the putative '*Ca*. L. asiaticus' contigs, one each for 454, SOLiD, and Solexa

data, as well as a collective reference assembly with all data from each method. Prior to performing the reference assemblies, the 34 '*Ca*. L. asiaticus' contigs from GenBank were assembled in SeqMan Pro Ver. 7.2.1 (DNAstar, Lasergene), which reduced the number of contigs to 32.

In order to confirm the '*Ca*. L. asiaticus' matches seen in citrus phloem were not the result of other alphaproteobacteria present in citrus phloem, reference assemblies with the sequenced genomes of 5 alphaproteobacterial relatives were performed. The genomes used in this analysis were selected by performing a local BLAST search of the '*Ca.* L. asiaticus' contigs against a database of 40 fully sequenced alphaproteobacteria, including a representative from each fully sequenced genus on NCBI (Accession numbers NC_009484.1, NC_003062.2, NC_004842.2, NC_009937.1, NC_008783.1, NC_010581.1, NC_004463.1, NC_010742.1, NC_002696.2, NC_009952.1, NC_007354.1, NC_007722.1, NC_011365.1, NC_006677.1, NC_008343.1, NC_008358.1, NC_007802.1, NC_007626.1, NC_008347.1, NC_002678.2, NC_010172.1, NC_007798.1, NC_007964.1, NC_007794.1, NC_009668, NC_009488.1, NC_008686.1, NC_009719.1, NC_007205.1, NC_011144.1, NC_008380.1, NC_007494.1, NC_008435.1, NC_007643.1, NC_009881.1, NC_008209.1, NC_009511.1, NC_002978.6, NC_009720.1, and NC_006526.1). The 5 species that were most closely related to '*Ca.* Liberibacter' were chosen for this study. Those 5 genomes were *R. leguminosarum* bv. *viciae* 3841 (NC_008378, NC_008379, NC_008380, NC_008381, NC_008382, NC_008383, & NC_008384), *Bartonella bacilliformis* KC583 (NC_008783.1), *Caulobacter crescentus CB15* (NC_002696.2), *Brucella abortus S19* (NC_010740 & NC_010742), and *Orientia tsutsugamushi* str. *Boryong* (NC_009488.1). '*Ca.* P. asteri' and *P. acnes* have also been reported to be associated with HLB infection in citrus; therefore, reference assemblies were performed with fully sequenced strains

of these species, '*Ca.* P. asteris' AYWB (NC_007716, NC_007717, NC_007718, NC_007719, NC_007720) and *P. acnes* KPA171202 (NC_006085). An additional reference assembly was performed against the genome of the non-HLB-associated bacterium, *M. luteus* NCTC 2665 (NC_012803), as a comparison.

To determine the detection limit for identifying a bacterial genome present in the metagenomic data, the mean and standard deviations for the fold coverage of *R. leguminosarum* bv. *viciae* 3841, *B. bacilliformis* KC583, *C. crescentus CB15*, *B. abortus S19*, *O. tsutsugamushi* str. *Boryong*, '*Ca.* P. asteris' AYWB, and *P. acnes* KPA171202 were calculated. As 99% of observed genomes are likely to fall within 3 standard deviations of the mean assuming a normal distribution, any genome with enough matching reads and bases to have a fold coverage greater than 3 standard deviations above the mean was considered significant. The detection limit for identifying a true match from the RDP database was determined using the same method by calculating the mean and standard deviation of the fold coverage from each match in the database.

With such a large amount of sequence data from citrus phloem, it is expected that the dataset also contains a high fold coverage of the citrus genome. To confirm this, a reference assembly of all citrus phloem data was performed against a citrus EST database. The citrus EST database was composed of all *C. sinensis* cv. Valencia ESTs from KEGG (http://www.genome.jp/kegg-bin/show_organism?org=ecsi). These ESTs were run in a de novo assembly on CLC Genomics Workbench to eliminate duplicates. Assembled contigs and non-assembled ESTs were used as the reference in the assembly with the complete citrus phloem metagenome dataset.

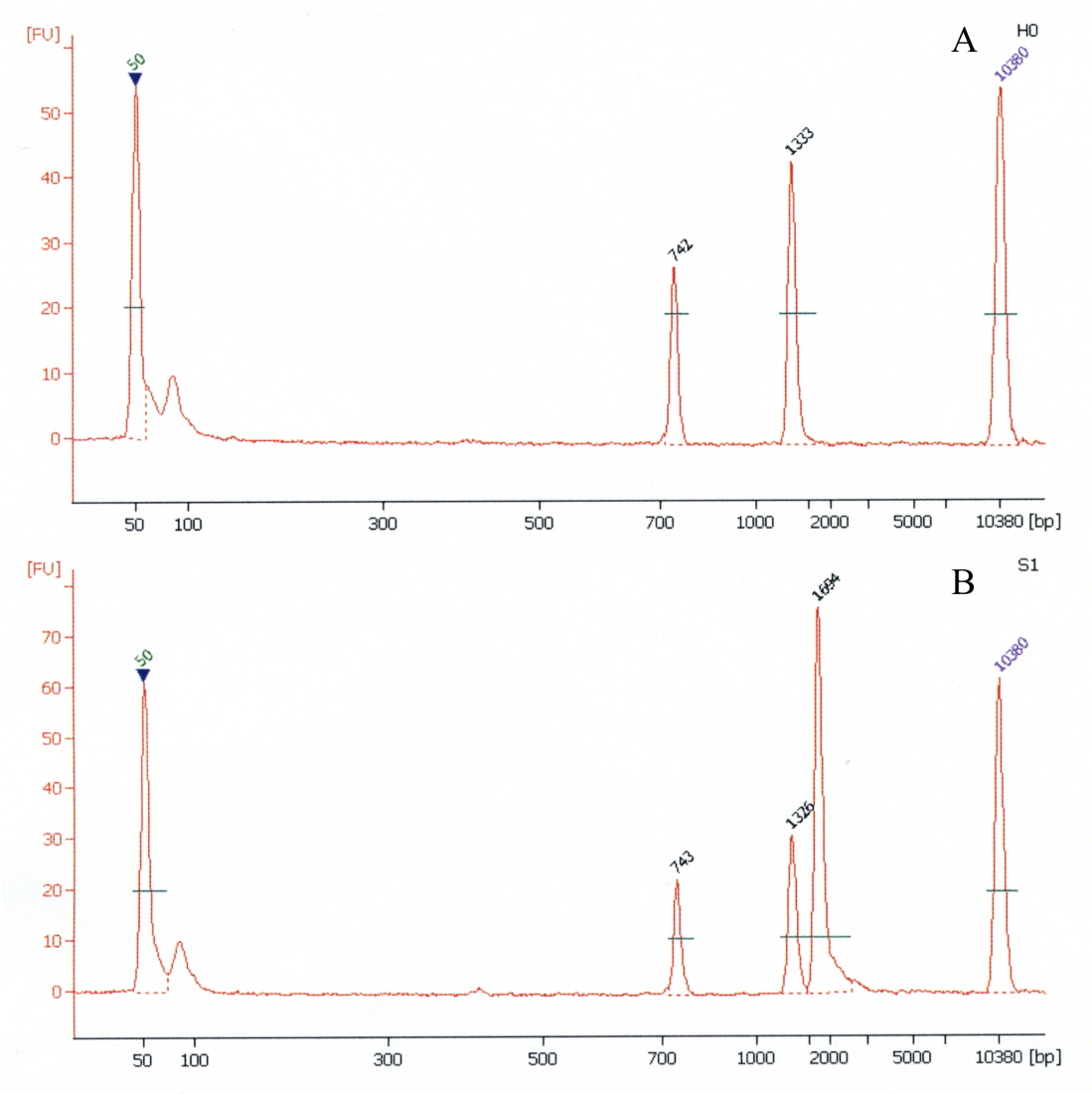


Figure 3-1. Automated ribosomal intergenic spacer analysis (ARISA) of healthy and Huanglongbing infected *Citrus sinensis* phloem. Example of the ARISA profiles for the DNA isolated from the phloem tissue of citrus samples, A) non-symptomatic and B) symptomatic, generated by the Agilent 2100 bioanalyzer. Base pair sizes are indicated adjacent to the peaks. The left-most (50 bp) and the right-most (10,380 bp) peaks represent the markers used to align the ladder data with data from the sample wells. The peaks of approximately 740 bp represent sequences amplified from the *C. sinensis* chloroplast, the peaks of approximately 1,330 bp represents the sequences amplified from the mitochondrial gene for mitochondrial RNA, and the peak of approximately 1,600 bp corresponds of the sequences amplified from '*Candidatus* Liberibacter asiaticus'.

Table 3-1. Overview of metagenomic sequence statistics

Sequencing technique	Number of reads	Number of gigabases	Average read length
454	246,355	0.058	238.48
Solexa	49,423,731	1.8	36
SOLiD	337,075,448	11.8	35
All data	386,745,534	13.6	35.26

Table 3-2. Fold coverage of each '*Candidatus* Liberibacter asiaticus' contig determined by reference assembly with 454, Solexa, and SOLiD metagenomic data from Huanglongbing-infected citrus phloem

Reference sequences[a]	454	Solexa	SOLiD	All datasets
NZ_ABQW01000001	0.054	14.314	12.221	26.595
NZ_ABQW01000002	0.056	15.028	12.005	27.089
NZ_ABQW01000003	0.08	13.949	9.146	23.175
NZ_ABQW01000004	0.083	15.048	13.107	28.272
NZ_ABQW01000005_33[b]	0.057	15.414	11.843	27.314
NZ_ABQW01000006	0.081	14.459	10.881	25.421
NZ_ABQW01000007	0.061	16.453	12.288	28.802
NZ_ABQW01000008	0.074	13.405	9.257	22.736
NZ_ABQW01000009	0.101	13.624	11.64	25.348
NZ_ABQW01000010	0.026	13.835	8.996	22.855
NZ_ABQW01000011	0.055	12.495	15.286	27.789
NZ_ABQW01000012	0.033	13.196	9.464	22.694
NZ_ABQW01000013	0.046	14.242	11.206	25.494
NZ_ABQW01000014	0.083	13.883	9.675	23.641
NZ_ABQW01000015	0.049	14.103	8.957	23.109
NZ_ABQW01000016	0.07	19.019	9.921	29.009
NZ_ABQW01000017	0.064	14.895	10.927	25.886
NZ_ABQW01000018	0.112	10.782	16.71	27.599
NZ_ABQW01000019	0.124	14.705	9.243	24.071
NZ_ABQW01000020	0.197	14.564	15.644	30.404
NZ_ABQW01000021	0.055	13.022	11.739	24.816
NZ_ABQW01000022_31[c]	0.341	22.494	29.517	52.363
NZ_ABQW01000023	0.164	16.9	20.3	37.364
NZ_ABQW01000024	0.088	11.734	13.978	25.8
NZ_ABQW01000025	0.099	26.224	14.892	41.211
NZ_ABQW01000026	0.219	10.862	18.077	28.904
NZ_ABQW01000027	0.106	15.764	17.352	33.209
NZ_ABQW01000028	0	15.026	5.616	20.642
NZ_ABQW01000029	0	14.958	11.047	26.042
NZ_ABQW01000030	0	13.966	8.51	22.477
NZ_ABQW01000032	0.333	28.444	31.289	60.067
NZ_ABQW01000034	0	15.95	12.042	27.913

[a]Each contig represented by its GenBank Ref Seq number. [b]Contig sequence assembled from NZ_ABQW01000005 and NZ_ABQW01000033 using SeqMan Pro. [c]Contig sequence assembled from NZ_ABQW01000022 and NZ_ABQW01000031 using SeqMan Pro.

Table 3-3. Percent length of each '*Candidatus* Liberibacter asiaticus' contig covered by 454, Solexa, and SOLiD metagenomic data from Huanglongbing-infected citrus phloem as determined by reference assembly

Reference sequences[a]	454	Solexa	SOLiD	All datasets
NZ_ABQW01000001	4.951	99.947	94.585	99.989
NZ_ABQW01000002	4.966	99.973	94.272	99.994
NZ_ABQW01000003	6.826	99.93	91.246	99.966
NZ_ABQW01000004	7.413	99.671	93.607	99.784
NZ_ABQW01000005_33[b]	5.022	99.936	92.817	99.947
NZ_ABQW01000006	7.205	99.991	92.966	99.999
NZ_ABQW01000007	4.917	99.927	92.189	99.946
NZ_ABQW01000008	6.635	99.97	91.62	99.995
NZ_ABQW01000009	7.100	99.891	93.702	99.997
NZ_ABQW01000010	2.545	99.928	91.094	99.964
NZ_ABQW01000011	3.075	99.393	95.527	99.691
NZ_ABQW01000012	3.295	99.927	93.125	100.002
NZ_ABQW01000013	3.899	99.964	93.291	99.964
NZ_ABQW01000014	3.223	99.952	91.694	99.952
NZ_ABQW01000015	4.942	99.964	91.593	100
NZ_ABQW01000016	6.984	99.991	91.402	99.995
NZ_ABQW01000017	5.456	99.989	93.582	99.989
NZ_ABQW01000018	7.112	99.469	98.340	100
NZ_ABQW01000019	10.086	99.983	93.184	99.983
NZ_ABQW01000020	16.012	99.939	97.128	99.983
NZ_ABQW01000021	4.232	99.893	94.603	100
NZ_ABQW01000022_31[c]	28.651	99.525	97.375	99.835
NZ_ABQW01000023	14.74	99.88	97.985	100
NZ_ABQW01000024	4.783	100	97.285	100
NZ_ABQW01000025	9.843	99.837	91.628	99.837
NZ_ABQW01000026	21.144	98	95.869	99.951
NZ_ABQW01000027	10.614	100	99.321	100
NZ_ABQW01000028	0	99.854	82.389	99.854
NZ_ABQW01000029	0	99.849	96.939	100
NZ_ABQW01000030	0	99.873	90.361	100
NZ_ABQW01000032	25.874	99.563	100	100
NZ_ABQW01000034	0	100	96.902	100

[a]Each contig represented by its GenBank Ref Seq number. [b]Contig sequence assembled from NZ_ABQW01000005 and NZ_ABQW01000033 using SeqMan Pro. [c]Contig sequence assembled from NZ_ABQW01000022 and NZ_ABQW01000031 using SeqMan Pro.

Table 3-4. All metagenomic phloem sequences compared to the contigs of '*Candidatus* Liberibacter asiaticus' or the fully sequenced genomes of other alphaproteobacteria by reference assembly

Bacterial species	Genome size (bp)	Number of matching reads	Number of matching bases	Fold coverage
'*Candidatus* Liberibacter asiaticus' str. psy62	1,217,424	900,124	32,071,418	26.34
Rhizobium leguminosarum bv. *viciae* 3841	5,892,908	8,027	281,989	0.048
Bartonella bacilliformis KC583	1,445,021	7,209	253,180	0.175
Caulobacter crescentus CB15	4,016,947	5,969	209,512	0.052
Brucella abortus S19	2,238,636	7,505	263,576	0.118
Orientia tsutsugamushi str. *Boryong*	2,127,051	2,339	82,333	0.039

Table 3-5. Fold coverage of the top 16S rRNA RDP database matches with 454, Solexa, and SOLiD metagenomic data from Huanglongbing-infected citrus phloem

Reference sequences	454	Solexa	SOLiD	All Data Sets
'*Candidatus* Liberibacter'	0.328	5.23	17.365	23.173
Nitrospina gracilis	0	0	1.427	1.451
'*Candidatus* Burkholderia'	0	0.375	0.885	1.301
Syntrophus gentianae	0	0.046	0.603	1.028
Nitratifractor salsuginis	0	0	0.942	0.991
Syntrophus buswellii	0	0.022	0.51	0.895
Pseudomonas rhizosphaerae	0	0.023	0.616	0.886
Hoeflea marina	0	0.073	1.229	0.85
Thermonema rossianum	0	0	0.766	0.837
Thermodesulfovibrio islandicus	0	0.231	0.645	0.804
Leucothrix mucor	0	0.094	0.685	0.802
Desulfobacter halotolerans	0	0.073	0.448	0.732
Pseudomonas argentinensis	0	0	1.049	0.685
Desulfospira joergensenii	0	0.048	0.506	0.627
Martelella mediterranea	0	0	0.993	0.615
Marinomonas aquimarina	0	0	0.509	0.554
Mycoplasma cloacale	0	0.025	0.427	0.527
Phaeospirillum fulvum	0	0.051	0.448	0.525
Desulfotignum phosphitoxidans	0	0.074	0.508	0.508
Desulfomonile tiedjei	0	0.14	0.369	0.508
Desulfobacula toluolica	0.03	0	0.223	0.345

Table 3-6. Comparison of '*Candidatus* Liberibacter' 16S rRNA to 454, Solexa, and SOLiD metagenomic sequences from Huanglongbing-infected phloem by reference assembly

Sequencing technique	Number of matching reads	Fold coverage	Length of consensus (bp)	% 16S gene matching consensus sequence
454	2	0.328	379	26.17
Solexa	211	5.23	950	65.61
SOLiD	720	17.37	1,118	77.21
All Data	944	23.17	1,212	83.7

Table 3-7. Comparison of '*Candidatus* Liberibacter asiaticus' and other reported Huanglongbing-associated bacteria to all metagenomic phloem sequences

Bacterial species	Genome size (bp)	Number of matching reads	Number of matching bases	Fold coverage
'*Candidatus* Liberibacter asiaticus' str. psy62	1,217,424	900,124	32,071,418	26.34
'*Candidatus* Phytoplasma asteris'	723,970	3,926	138,117	0.191
Propionibacterium acnes KPA171202	2,560,265	1,744	61,452	0.024
Micrococcus luteus NCTC 2665	2,501,097	2,677	94,125	0.038

Table 3-8. Comparison of metagenomic assemblies with total metagenomic data to metagenomic data with chloroplast sequences removed

Bacterial species	Total fold coverage	Fold coverage minus chloroplast	Total bases matched	Bases matched minus chloroplast
'*Candidatus* Liberibacter asiaticus' str. psy62	26.34	26.179	32,071,418	31,871,539
Rhizobium leguminosarum bv. *viciae* 3841	0.048	0.018	281,989	104,598
Bartonella bacilliformis KC583	0.175	0.058	253,180	83,988
Caulobacter crescentus CB15	0.052	0.014	209,512	57,130
Brucella abortus S19	0.118	0.038	263,576	84,194
Orientia tsutsugamushi str. *Boryong*	0.039	0.017	82,333	35,218
'*Candidatus* Phytoplasma asteris'	0.191	0.016	138,117	11,356
Propionibacterium acnes KPA171202	0.024	0.009	61,452	22,352
Micrococcus luteus NCTC 2665	0.038	0.004	94,125	9,822

Table 3-9. '*Candidatus* Liberibacter asiaticus' contig primers

Primer name	Sequence (5' – 3')
NZ_ABQW01000001F	CTGGTGTTCGCAAGGGTATT
NZ_ABQW01000001R	ACACCCGAACGCATTAACTC
NZ_ABQW01000002F	ACTGCCCTGCTTTTTCTTCA
NZ_ABQW01000002R	ACAGGTCGTGCTATGGTTCC
NZ_ABQW01000003F	TGGAAATTTCGGAGAGGATG
NZ_ABQW01000003R	GAAATGTTGCTGCGACTTCA
NZ_ABQW01000004F	TGATTCTCCTCATCCCCAAG
NZ_ABQW01000004R	CTTTGGTGGTATTCGCTGGT
NZ_ABQW01000006F	GATTTAGCGGAATCGGATCA
NZ_ABQW01000006R	TCATGCCTTGCTTGAAACAG
NZ_ABQW01000007F	CGCATCCACAGTTTATGGTG
NZ_ABQW01000007R	AGCCGCTACAGCAATCCTAA
NZ_ABQW01000008F	GGTTCCAGAAGTTGCTGCTC
NZ_ABQW01000008R	CAACACGCCCCTTATCTTGT
NZ_ABQW01000009F	ACGAGCGTAACCAACAATCC
NZ_ABQW01000009R	AGGCTCTTGGACTCACGAAA
NZ_ABQW01000010F	CTGCTTCAATAAGCGGGAAG
NZ_ABQW01000010R	CTGAGCAACATGGCTCAAAA
NZ_ABQW01000011F	CTGTTACGGTTGCTCCCATT
NZ_ABQW01000011R	CGAATTGTCTAGCAGCCACA
NZ_ABQW01000012F	GGGCGTACAACCCAACTCTA
NZ_ABQW01000012R	TTTAACGAGCCGCTTCAAGT
NZ_ABQW01000013F	TACGCCTGCATATGCTTCAG
NZ_ABQW01000013R	TGTCATGAAGCGAGAAAACG
NZ_ABQW01000014F	ACCCAAAAGTCGGTGTCTTG
NZ_ABQW01000014R	CCGTAAACACGAAAGGCAAT
NZ_ABQW01000015F	GGGGCAAGCTCATGATAAAA
NZ_ABQW01000015R	AGTTGCGAGAAGGCTTGGTA
NZ_ABQW01000016F	CTTGATCTCGCCACTGTTGA
NZ_ABQW01000016R	CATTCGCAGTGCTATGGCTA
NZ_ABQW01000017F	ACGACAGGGGTGCAAGTATC
NZ_ABQW01000017R	TGCATCAGAAAGCGCAATAC
NZ_ABQW01000018F	ACGGGGCTTGGTAAAGAAGT
NZ_ABQW01000018R	CAAACTCCCGTTTCACACCT
NZ_ABQW01000019F	AGAGTCATGGAAGCGGCTAA
NZ_ABQW01000019R	CCGAGGGATATCAGCGATTA
NZ_ABQW01000020F	GAAGCAATATGGGCTGGAAA
NZ_ABQW01000020R	TTCAGAAGACATGCCACGAG
NZ_ABQW01000021F	AATTGGAACTTCCCCCAAAC
NZ_ABQW01000021R	ACGATGCTTTGCTCGATCTT
NZ_ABQW01000023F	CCCTCTTCCAATCCCTTCTC
NZ_ABQW01000023R	CTATATCGGCAGGTCGTGGT
NZ_ABQW01000024F	CTCTTGCATCCAACGACAGA

Table 3-9. Continued

Primer name	Sequence (5’ – 3’)
NZ_ABQW01000024R	GGACGGTTGTCTTACCCTCA
NZ_ABQW01000025F	GGGAGTACGGAGAGTGGTGA
NZ_ABQW01000025R	TTGATAAAGACGCGCTGTTG
NZ_ABQW01000026F	CGGCGAAAGAGCTTTACAAC
NZ_ABQW01000026R	CTGTGGCAATAGCGTTGAGA
NZ_ABQW01000027F	CGAGAAGGGGGAGGATTAAG
NZ_ABQW01000027R	AGCGTTATAGGCTCCAGCAA
NZ_ABQW01000028-2kF	TGCAATCAAGGCTCACCATA
NZ_ABQW01000028-2kR	TCCTACTCCTTCCGCCATAA
NZ_ABQW01000029-2kF	CAATGAGAGATAAGACAATCCTGTTT
NZ_ABQW01000029-2kR	GCTTAAAGCTGATAAACCCATCA
NZ_ABQW01000030F-1kF	TCCAGAATCCCGTCTTGTTC
NZ_ABQW01000030R-1kR	GTGGCTTTGGGTTTGTCTGT
NZ_ABQW01000032F-1kF	GGGCATTGAGCCTGAATCTA
NZ_ABQW01000032R-1kR	TGCAAAACGAGAAACTGACG
NZ_ABQW01000034F-1kF	CTTGTCGGCAAGTTGGTTTC
NZ_ABQW01000034R-1kR	ATGGCGAAAAACGACAGTTC
NZ_ABQW01000022_31F	GGAACAAAGCAAACGAGAGC
NZ_ABQW01000022_31R	CAAAAGGAGGAAGTCGCAAG
NZ_ABQW01000033_05F	GTTGTGTGCCGAGGATAGGT
NZ_ABQW01000033_05R	GAGGACAGACCGCTCTCAAC

CHAPTER 4
OPTICAL MAPPING OF *GLUCONACETOBACTER DIAZOTROPHICUS* PAL 5 REVEALS CHROMOSOMAL REARRANGEMENTS IN COMPLETED GENOME SEQUENCE

Background

Gluconacetobacter diazotrophicus PAl 5 is a bacterial endophyte of sugarcane, originally isolated in Brazil (Gillis et al. 1989). This endophyte is of agricultural significance due to its ability to provide fixed nitrogen to its host plant, in addition to increasing plant growth by mechanisms independent of nitrogen fixation (Lee et al. 2004; Sevilla et al. 2001). The ability of *G. diazotrophicus* to increase growth and reduce plant dependence on nitrogen fertilization also makes it important to increasing the efficiency of biofuel production from sugarcane (Boddey 1995). Since it was first isolated, other strains of *G. diazotrophicus* have been isolated in several other countries and plant hosts (Boddey et al. 1995; Fuentes-Ramirez et al. 1993; Hoefsloot et al. 2005; Paula et al. 1991; Jimenez-Salgado et al. 1997; Tapia-Hernandez et al. 2000). As a result, there has been great interest in sequencing the genome of *G. diazotrophicus* to guide further research on this bacterium and to better understand endophytic nitrogen fixation by comparative genomics with other sequenced nitrogen fixing endophytic bacteria.

The genomic sequencing of *G. diazotrophicus* PAl 5 was undertaken by two independent groups, RioGene in Brazil, funded by FAPRJ, and the United States Department of Energy's Joint Genome Institute (JGI) in California. Both groups have since closed the genomic sequence of *G. diazotrophicus* PAl 5 and deposited it on GenBank (Accession numbers CP001189.1 and AM889285.1). Interestingly, though both groups reported sequencing the PAl 5 strain, the two genome sequences varied greatly between each other in gene arrangement and plasmid content, indicating the presence of sequencing and/or assembly errors in one of the genomic sequences. To ascertain which genome assembly more closely matches the physical map of PAl 5, optical mapping was utilized.

Optical mapping serves to create a physical restriction map of a genome assembled from deoxyribonucleic acid (DNA) molecules immobilized on a glass slide prior to digestion with a selected restriction enzyme, maintaining the original order of restriction fragments. After digestion, DNA is stained and visualized by fluorescent microscopy, and the resulting digitized images are analyzed in an assembly program to construct an optical restriction map of the genome of interest (Aston et al. 1999; Zhou et al. 2004a). These optical maps can be compared to *in silico* digests of DNA sequences and have been utilized in many sequencing studies, serving as scaffolds for contig alignment, as well as an independent means of identifying errors (inversions, insertions, deletions, translocations, etc.) in previously assembled sequences (Latreille et al. 2007; Lim et al. 2001; Reslewic et al. 2005; Wu et al. 2009; Zhou et al. 2007; Zhou et al. 2002; Zhou et al. 2004b). Therefore, optical mapping was deemed to be an ideal tool to elucidate which PAl 5 genome sequence most closely matched the physical DNA.

Results

Optical Map of *G. diazotrophicus* PAl 5

A BglII optical map of *G. diazotrophicus* PAl 5 (ATCC 49037) was constructed in order to determine which genome assembly was the most accurate representation of the strain. The optical map was 3,845,512 basepairs (bp) in length and composed of 424 restriction fragments, with an average fragment size of 9,070 bp (Table 4-1). In comparison, the *in silico* map of the JGI sequence was 3,887,492 bp in length, while the RioGene map was 3,944,163 bp (Table 4-1). The average fragment length of both *in silico* maps is over 1,000 bp shorter than the average fragment length of the optical map (Table 4-1). These differences between the optical and *in silico* maps are likely due to the fact that restriction fragments shorter than 500 bp are not detected by optical mapping owing to such short fragments being washed off the optical slide (Meng et al. 1995).

Identification of Sequence Rearrangements Using Optical Mapping

Once the BglII optical map of PAl 5 was aligned to *in silico* BglII restriction maps generated from the two separate genome sequences, it was readily apparent that the sequence from RioGene contained numerous chromosomal rearrangements (Figure 4-1a). Comparison of the optical map to the RioGene *in silico* map revealed the presence of 2 large inverted regions (Figure 4-1b). These inversions were 555.9 and 564.3 kilobases (kb) in length, together spanning close to 28% of the genome sequence (Table 4-2). In addition, numerous translocations were identified in the RioGene sequence. One large translocation spanning 865.8 kb of the genome (Figure 4-1c) and 5 smaller translocations ranging in size from 69.8 to 330.8 kb (Figure 4-1d) were identified (Table 4-2). From these determinations, it appears that 74% of the PAl 5 genome sequence proposed by RioGene is rearranged compared to the physical map of the PAl 5 genome. In contrast, the *in silico* map of the JGI PAl 5 sequence showed a higher alignment to the optical map (Figure 4-1a). Only 3 small inversions and 1 small translocation were detected. These regions only accounted for 5.6% of the genomic sequence.

Additionally, several regions of the RioGene BglII *in silico* map did not align to the optical map (Figure 4-1a). Together, these regions totaled 1,053,347 bp, or 26.7% of the genome sequence (Table 4-3). In comparison, the regions of the JGI *in silico* map that did not align to the optical map were composed largely of single restriction fragments, most of which were 500 bp in length or less (Table 4-3), which is the below the detection threshold of optical mapping technology. All told, it appears that nearly all of the PAl 5 genome sequence from RioGene was either rearranged or did not align to the optical map of PAl 5. From these analyses, it was concluded that the sequence from JGI is a more accurate representation of PAl 5's genome.

Differences in Annotation Between Genome Sequences

Given the surprisingly high level of chromosomal rearrangements and non-aligned regions between genomic sequences reported from the same strain, the annotations of both PAl 5 sequences were determined using the Rapid Annotation using Subsystem Technology (RAST) (Aziz et al. 2008) web based annotation service to ascertain what effect these rearrangements have on gene calling. With a total of 8 rearrangements in the RioGene sequence, there are up to 16 locations where coding sequences (CDSs) could have been disrupted. Interestingly, 168 and 187 of the CDSs identified in the RioGene and JGI genomes, respectively, were unique, sharing 0% identity with CDSs in the other genome (Table 4-4). In total, 247 of the CDSs in the RioGene sequence shared less than 50% identity with CDSs in the JGI sequence (Table 4-4). This number of differences between the two genome sequences was over 10 times greater than expected from the observed inversions and translocations in the RioGene sequence. Since both genomic sequences are from the same strain, a similar complement of genes was expected. However, only 90% of the CDSs predicted in each genome shared greater than 90% identity (Table 4-4). Considering both genome sequences were generated from the same strain, the level of differences between them at the sequence level was surprising.

Fewer differences were seen between the two genomes when looking at the functional roles genes were assigned to. In total, 13 and 21 functional roles identified were found to be unique to the RioGene and JGI sequences, respectively (Table 4-5). Interestingly, two of the unique functional roles identified in the RioGene sequence were C- and N-terminal sections of a transketolase enzyme, though complete transketolase genes were identified in both the genome sequences. While this observation initially suggested a rearrangement had occurred within a transketolase gene in the RioGene sequence, examination of the genome region containing these transketolase genes showed that the C-terminal gene was immediately upstream of the

N-terminal section. This configuration of transketolase genes was also observed in 3 other bacteria in the SEED database (*Ruegeria* sp. PR1b, *Desulforubris audaxviator*, and *Carboxydothermus hydrogenoformans*), indicating that this split is not the result of a chromosomal rearrangement. In confirmation of this, none of the inversion or translocation break points fell between these two genes.

Given the number of inversions and translocations in the RioGene sequence, the annotation was also checked for transposases that could contribute to chromosomal rearrangements. The RioGene PAl 5 sequence was found to possess 110 transposase genes, while the JGI sequence only contained 59 transposase genes, almost half that amount. A large number of these were putative transposases, though several insertion sequence (IS)3, IS4, and IS5 family proteins were also identified as well (Table 4-6). Interestingly, only 2 translocated regions had transposases less than 10 kb from each of their ends. A 69.7 kb translocation had an IS4 transposase family protein 6.9 kb upstream and an IS5 transposase family protein 814 bp downstream of the region. Another 171 kb rearrangement had putative transposases 193 bp and 6.9 kb up and downstream of the translocated region. A third translocated region had a putative transposase 348 bp upstream, but the nearest downstream transposase was greater than 30 kb away. In all other cases, the nearest transposase to an inverted or translocated region was over 18 kb away.

Discussion

The production of two different genome sequences from the same bacterium, *G. diazotrophicus* PAl 5, demonstrated the need to confirm the assembly of these genomes through an independent method. Optical restriction mapping has been used by many groups as an independent method of verifying sequences and identifying assembly errors, due to the fact that it maintains the order of restriction fragments in the mapping process. To that end, these two genome assemblies were compared to a BglII optical map of the PAl 5 bacterium, leading to the

determination that the sequence reported by JGI is a more accurate representation of the PAl 5 strain from the American Type Culture Collection (ATCC) while the sequence reported by RioGene contained numerous rearrangements, including 2 large inversions and several translocations, when compared to the physical map of the PAl 5 strain.

In the current study, optical mapping was used to distinguish between discordant genomic assemblies of the same bacterial strain. The size and number of chromosomal rearrangements identified in the RioGene sequence of *G. diazotrophicus* PAl 5 was high, with nearly all of the sequence composed of regions that were inverted, translocated, or not aligned to the optical map of the PAl 5 strain. In contrast, only a few small inversions were detected when the JGI PAl 5 sequence was compared to the optical map. In addition, annotation of the two genome sequences found that approximately 5% of the CDSs in each were unique to each genome sequence. This is a surprisingly high amount considering the two genomes are reported to be from the same strain and much greater than would be expected from the observed sequence rearrangements. There are a few possibilities for the differences between these two PAl 5 genome sequences.

One explanation for the differences between the RioGene and JGI sequences is assembly errors. While many studies have reported using optical maps to aid in genome assembly and identification of assembly errors prior to completion, fewer have reported using this technique to identify errors in previously completed genomes. After the successful use of optical mapping to aid in assembling the *Xenorhabdus nematophila* genome, Latreille and associates (2004) used the same technique on another *Xenorhabdus* species, *X. bovienii*, that had been previously sequenced, identifying a large inversion in the genome assembly that had been considered finished. In addition, optical mapping has also been used to verify assemblies between strains of the same species. In the case of *Mycobacterium avium* subspecies *paratuberculosis*, an optical

map of the ATCC type strain was used to reveal the presence of an inversion in the genome of the sequenced strain, which was determined to be due to an assembly error rather than genomic variation between strains (Wu et al. 2009). These two instances illustrate how even "closed" and published genomes may contain significant assembly errors, indicating that caution should be taken when looking at assemblies where optical mapping was not used.

Another reason for the differences between the RioGene and PAl 5 sequences is natural divergence and evolution that can occur during culturing, though the extremely high level of differences between the two sequences indicates other factors may be involved. The amount of sequence rearrangement seen between the PAl 5 optical map and the RioGene PAl 5 sequence is higher than the level of rearrangements seen between different strains within the same species. Comparison of an optical map of *E. coli* H10407 to the sequence of *E. coli* K-12 showed no major structural differences between the two strains (Chen et al. 2006). In *M. avium* subspecies *paratuberculosis*, only 1 inversion between the sequenced strain, K-10, and the optically mapped strain, ATCC 19698, was detected and that inversion was subsequently determined to be due to an assembly error rather than a true chromosomal rearrangement (Wu et al. 2009). In addition, when comparing *Shigella flexneri* strains 2457T and 301, Zhou and associates (2004a) found 3 inversions that were 876, 72, and 20 kb in length. Given these numbers, if the sequence inversions and translocations seen between the RioGene PAl 5 sequence and the PAl 5 optical map are due to true chromosomal rearrangements that occurred as a result of evolution during culturing, one must question whether they can still be considered the same strain.

If the breakpoints of assembly errors such as inversions or translocations occur within a coding region, such errors could alter the annotation of the genome. For example, when the previously mentioned inversion in the sequence of *M. avium* subspecies *paratuberculosis* K-10

was corrected, 2 new genes were identified (Wu et al. 2009). As a result, the annotation of the PAl 5 sequences from both RioGene and JGI were determined and compared using the RAST on-line annotation pipeline (Aziz et al. 2008). Both sequences were reannotated instead of using the original annotation provided on GenBank to avoid any biases based on differences in annotation methods. Six percent of the CDSs from each genome shared less than 50% identity when compared against each other. Interestingly, approximately 5% of the CDSs showed 0% identity when the two genomes were compared, equating to 168 and 187 unique genes in each sequence. Again, this level of sequence and gene difference was surprising from what is supposed to be the same strain, even with chromosomal rearrangements. Possible reasons for such differences could be evolution of the strain, or sequencing errors. Annotation of both genomes also revealed that the RioGene sequence possessed almost twice as many transposases as the JGI sequence. The strikingly high number of transposases in the RioGene sequence in relation to the JGI sequence suggests the possibility that some of the sequence rearrangements seen may be the result of transposition. Alternatively, since 16 of the transposases originated from IS sequences, which are flanked by inverted repeats (Mahillon and Chandler 1998), it is also possible that these repeated regions caused errors in assembly.

The *G. diazotrophicus* PAl 5 optical map was constructed from a PAl 5 isolate obtained directly from the ATCC and maintained in -80^{o}C freezer stocks until cultured to obtain DNA for the optical map. Therefore, it can be said with confidence that the PAl 5 optical map is an accurate representation of the original PAl 5 strain submitted to ATCC. If the observed chromosomal inversions and translocations are the result of evolution in culture, this would indicate that the bacterium RioGene sequenced has evolved to the point that it may no longer be considered the same as the original PAl 5 strain. If sequencing errors are at fault, only

examination of the raw sequencing data could identify such a cause. The true origin of the differences between these two sequences, whether it be genuine chromosomal rearrangements that occurred during culturing of the bacteria or a combination of errors in sequencing and assembly, is difficult to ascertain without examining the raw sequencing reads or the specific isolate used by RioGene, which were unavailable. Given the degree of rearrangements seen between these two sequences, it is possible a combination of chromosomal rearrangement and assembly errors may be the cause.

Conclusions

The genomic sequence of *G. diazotrophicus* PAl 5 produced by JGI was deemed to be the most accurate representation of the genome as determined by optical mapping, while the RioGene sequence contained numerous rearrangements. Therefore, subsequent studies of this bacterium involving examination of the genome should utilize the JGI sequence. The observations made here further confirm the utility of optical mapping in determining proper assembly of genomic sequences and identifying potential chromosomal rearrangements. It also highlights the need to provide raw reads and quality scores when submitting genomes to allow for independent confirmation of assembly. As technology advances and new and improved annotation programs are developed, data from instances where contradictory sequences are observed could be reanalyzed in order to clarify results.

The rearrangements in the RioGene sequence of *G. diazotrophicus* PAl 5 may not have been identified had JGI not released a conflicting genome sequence of the same strain that prompted further investigation. As the genome sequencing of a single bacterial strain is not usually performed independently by different groups, the possibility remains that other previously released, closed genomes could contain similar differences compared to other bacterial isolates under the same strain designation maintained in different laboratories. Such

rearrangements in genome sequences of the same strain could confound future work by researchers using comparative genomics to look for variations between closely related organisms. In such cases, again, the best tool to distinguish actual variations between organisms will be optical mapping. Therefore, it is proposed that raw sequencing reads with quality scores be made available when publishing completed genomes and that optical mapping become a regular tool in genome assembly projects to ensure that differences which arise between genomic sequences from the same or related strains are genuine chromosomal rearrangements as opposed to assembly or sequencing errors.

Methods

Bacterial Strain

The bacterial strain used in this study was *Gluconacetobacter diazotrophicus* PAl 5 obtained from the American Type Culture Collection (ATCC 49037). *G. diazotrophicus* PAl 5 was cultured on yeast mannitol agar (YMA) and broth at 30^{o}C.

Preparation of Cells for Optical Mapping

G. diazotrophicus PAl 5 was grown in a 5 mL YM broth until the cells reached a density of 10^9 colony forming units per mL. The culture was dispensed into five 1.5 mL microcentrifuge tubes in 1 mL aliquots. Tubes were then centrifuged at 6,000 rpm for 10 minutes to pellet the cells. Tubes with cell pellets were shipped on dry ice to OpGen Technologies, Inc. (Madison, Wisconsin) for optical mapping.

Optical Mapping and Analysis

A BglII optical map of *G. diazotrophicus* PAl 5 was constructed by OpGen Technologies, Inc. (Madison, Wisconsin). *In silico* BglII restriction maps of the two complete *G. diazotrophicus* PAl 5 genomic sequences on GenBank (Accession numbers CP001189 and AM889285) were constructed from each sequence's GenBank file and compared to the BglII

optical map of PAl 5 using MapViewer version 2.1.1 (OpGen Technologies, Inc.). Plasmid sequences associated with each genome assembly did not align to the optical map and were therefore not included the analysis.

Comparison of Annotation

The annotations of the two genomic assemblies were determined using RAST ver. 2.0 (Aziz et al. 2008). Genome and plasmid sequences for RioGene (Accession numbers AM889285, AM889286, and AM889287) and JGI (Accession numbers CP001189 and CP001190) were concatenated into single FASTA files prior to RAST analysis. Annotations determined by RAST were compared using the SEED viewer (ver. 2.0) (Overbeek et al. 2005) based on percent identity between CDSs and the functional roles assigned to annotated genes.

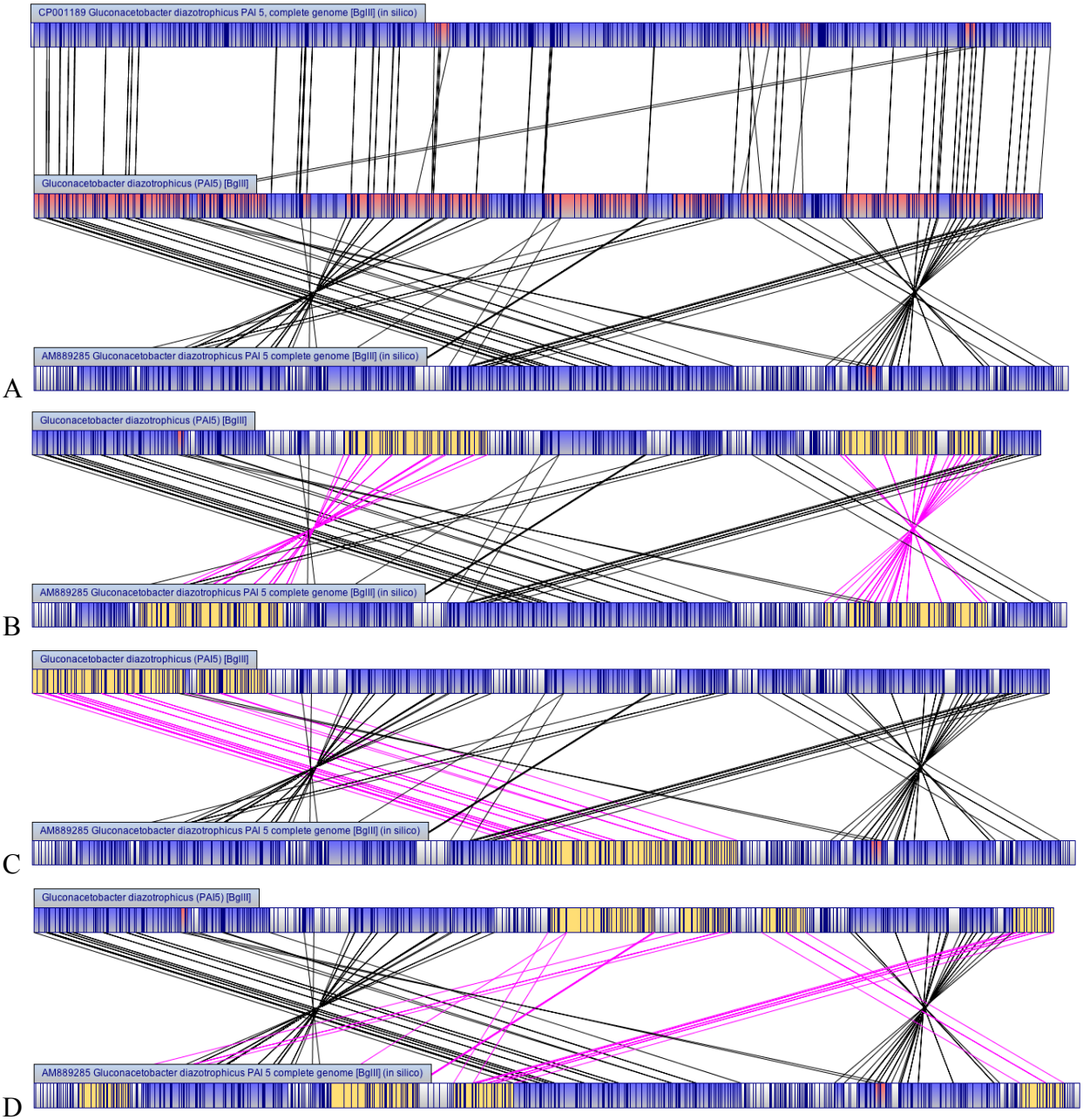


Figure 4-1. Alignment of *G. diazotrophicus* PAl 5 optical map with *in silico* maps of genome sequences. A) The BglII optical map of PAl 5 aligned against *in silico* optical maps calculated from the genome sequence proposed by RioGene (AM889285) and JGI (CP001189). B-D) Rearrangements in the RioGene PAl 5 sequence when aligned against the optical map. B) Two large inversions in RioGene sequence. C) Large translocation in RioGene sequence. D) Five translocations in RioGene sequence. Dark blue represents cut sites, light blue represents aligned regions, red represents regions aligning to both sequences, and white represents unaligned regions. Alignment lines for inversions and translocations are highlighted in pink. Inverted and translocated regions highlighted in yellow.

Table 4-1. Optical and *in silico* BglII restriction maps for *G. diazotrophicus* PAl 5

	Optical map	JGI *in silico* map	RioGene *in silico* map
Map length (bp)	3,845,512	3,887,492	3,944,163
Number of fragments	424	486	503
Average fragment length (bp)	9,070	7,999	7,841
Maximum fragment length (bp)	52,064	51,728	50,690
Minimum fragment length (bp)	562	24	28

Table 4-2. Rearrangement positions in *G. diazotrophicus* PAl 5 genome sequence from RioGene

Rearrangement type	Start position (bp)	Stop position (bp)	Length (bp)
Inversion	391,267	955,614	564,347
Inversion	3,078,324	3,634,241	555,917
Translocation	149,268	358,682	209,414
Translocation	1,115,930	1,446,765	330,835
Translocation	1,581,823	1,651,595	69,772
Translocation	1,627,253	1,796,352	169,099
Translocation	1,796,352	2,662,168	865,816
Translocation	3,706,896	3,878,171	171,275

Table 4-3. Regions of *in silico* maps not aligned to the *G. diazotrophicus* PAl 5 optical map

	JGI	RioGene
Total length of unaligned regions (bp)	27,540	1,053,347
Average unaligned fragment length (bp)	574	5,885
Maximum unaligned fragment length (bp)	1,341	32,719
Minimum unaligned fragment length (bp)	24	28

Table 4-4. Comparison of coding sequences between two genomic sequences of *G. diazotrophicus* PAl 5 based on percent identity

Percent identity to comparison genome	Number CDS from JGI	Percent CDS from JGI	Number of CDS from RioGene	Percent of CDS from RioGene
100%	2,024	56.7	2,069	56.0
≥99%	2,812	79.2	2,876	77.8
≥90%	3,190	89.8	3,313	89.6
>75%	3,267	92.0	3,402	92.0
≥50%	3,326	93.7	3,449	93.3
<50%	225	6.3	247	6.7
0%	187	5.3	168	4.5

Table 4-5. Unique functional roles between *G. diazotrophicus* PAl 5 genome sequences

Roles unique to JGI	Roles unique to RioGene
Ribose ABC transport system, periplasmic ribose-binding protein RbsB (TC 3.A.1.2.1)	Sorbitol dehydrogenase (EC 1.1.1.14)
D-alanine-D-alanine ligase (EC 6.3.2.4)	Transketolase, C-terminal section (EC 2.2.1.1)
UDP-N-acetylenolpyruvoylglucosamine reductase (EC 1.1.1.158)	Transketolase, N-terminal section (EC 2.2.1.1)
Organic hydroperoxide resistance protein	COG0028: Thiamine pyrophosphate-requiring enzymes
Organic hydroperoxide resistance transcriptional regulator	D-galactonate regulator, IclR family
Molybdenum cofactor biosynthesis protein B	Epi-inositol hydrolase (EC 3.7.1.-)
Flagellar biosynthesis protein fliL	Chromosome partition protein smc
Flagellar hook-associated protein flgL	dTDP-rhamnosyl transferase RfbF (EC 2.-.-.-)
Deoxyuridine 5'-triphosphate nucleotidohydrolase (EC 3.6.1.23)	Protein of unknown function DUF374
Aminopeptidase S (Leu, Val, Phe, Tyr preference) (EC 3.4.11.-)	Nicotinate-nucleotide adenylyltransferase (EC 2.7.7.18)
Leucyl/phenylalanyl-tRNA-protein transferase (EC 2.3.2.6)	DNA repair exonuclease family protein YhaO
Cysteinyl-tRNA synthetase (EC 6.1.1.16)	ATP-dependent DNA helicase UvrD/PcrA, proteobacterial paralog
tRNA:Cm32/Um32 methyltransferase	Outer membrane lipoprotein carrier protein LolA
	DNA-binding response regulator KdpE
	Osmosensitive K+ channel histidine kinase KdpD (EC 2.7.3.-)
	Potassium-transporting ATPase A chain (EC 3.6.3.12) (TC 3.A.3.7.1)
	Potassium-transporting ATPase B chain (EC 3.6.3.12) (TC 3.A.3.7.1)
	Beta-hexosaminidase (EC 3.2.1.52)
	Potassium-transporting ATPase C chain (EC 3.6.3.12) (TC 3.A.3.7.1)
	Protein-export membrane protein secD (TC 3.A.5.1.1)
	H^+/Cl^- exchange transporter ClcA

Table 4-6. Transposases in *G. diazotrophicus* PAl 5 genome sequences

	JGI	RioGene
Total transposase genes	59	110
Transposase	6	19
Transposase (class II)	1	2
Transposase (class III)	1	0
Transposase (class IV)	1	0
Putative transposase	27	64
Transposase IS3 family protein	2	4
Transposase IS3/IS911 family protein	1	0
Transposase IS4 family protein	6	4
Transposase IS5 family protein	4	7
Transposase IS256	1	0
Transposase IS630	0	1
Isrso16-transposase OrfA protein	1	0
Transposase and inactivated derivative	2	1
Transposase mutator type	5	6
Probable insertion sequence transposase protein	1	0
TRm2011-2a transposase	0	2

CHAPTER 5
ENDOPHYTE MEDIATED PLANT GROWTH PROMOTION

Introduction

It has been long known that some bacteria have beneficial effects on plant growth. Some of these bacteria are believed to enhance growth by providing fixed nitrogen to plants, but are also known to increase growth yield by mechanisms independent of nitrogen fixation. These bacteria can be rhizobacteria that live in the soil around plant roots, or endophytes, living inside the plant tissue without forming symbiotic structures or causing disease. One such endophytic bacteria, *Klebsiella pneumoniae* 342 (Kp342), was isolated from a nitrogen efficient line of maize (Chelius and Triplett 2000). Kp342 is a model for endophytic bacteria because it has been shown to colonize the interior of several plant species, including soybean (*Medicago sativa* and *M. truncatula*), wheat (*Triticum aestivum*), rice (*Oryza sativa*), and *Arabidopsis thaliana* (Dong et al. 2003a and b). In addition, Kp342 increases the yield of maize grown in the field and *Arabidopsis* grown in the greenhouse (Riggs et al. 2001 and 2002). These increases in plant growth were independent of nitrogen fixation since the plants tested were grown under conditions of nitrogen fertilization when nitrogen fixation would be inhibited (Riggs et al. 2001 and 2002). Kp342 has also been found to fix nitrogen within wheat. It has been shown to produce the nitrogenase enzyme within the cortex of maize and wheat, and inoculation with this bacterium relieves nitrogen deficiency symptoms in Trenton wheat (Chelius and Triplett 2000; Iniguez et al. 2004).

In the course of studying this bacterium, confocal microscopy showed that Kp343 colonization is localized around points of lateral root emergence, indicating that the bacteria may be entering plants through cracks formed as lateral roots emerge from the primary root (Dong et al. 2003a; Iniguez et al. 2004). During an experiment to determine if it enters plants in this

manner, Kp342 was discovered to increase the number of lateral roots formed by *Arabidopsis*. This unexpected result yielded an interesting development in the study of Kp342's nitrogen fixation independent growth promotion. First, the direct consequence of a larger, more branched root system is that plants can obtain more nutrients and water from the soil, enhancing the growth of the plant. Second, the increase in lateral root numbers provided a quick assay for future experiments to determine Kp342's growth-promoting mechanisms, since the signal responsible for lateral root promotion may also be involved in an overall increase in growth yield.

Lateral Root Development

Lateral roots develop from a group of differentiated, pericycle cells in contact with xylem pole cells called founder cells, and their development into lateral roots is regulated by auxin (Malamy and Benfey 1997; Casimiro et al. 2003). These pericycle cells are signaled to become lateral root primordia when they pass through the root elongation and differentiation zones, but lateral root primordium formation is also signaled by a second check point that is dependent upon environmental factors later in development (Dubrovsky et al. 2000). Lateral root formation begins when lateral root primordia develop from founder cells that undergo anticlinal divisions resulting in 2 short daughter cells surrounded by 2 long daughter cells (Stage I primordia) and then undergo periclinal divisions, creating a second layer of cells resulting in stage II primordia (Malamy and Benfey 1997). In addition to cell cycle activation, the fate of these pericycle cells must be respecified in order to undergo lateral root initiation (Vanneste et al. 2005). After initiation, cells in the lateral root primordia continue to divide and progress to stage VII and emerge from the primary root by cell expansion (Malamy and Benfey 1997). Upon emergence, the lateral root apical meristem is formed (Malamy and Benfey 1997). It is at this point, when

cells at the lateral root tip are short and more numerous, that the primordium is considered to have developed into a lateral root (Malamy and Benfey 1997).

The molecular events between auxin signaling and lateral root formation are not as well understood as the stages of morphological development, partially because the few cells involved in the process are embedded in the root and difficult to access (Himanen et al. 2004). To overcome this obstacle, Himanen and associates (2004) synchronously induced pericycle cells to form lateral roots, increasing the number of cells undergoing the process in order to detect changes in gene expression associated with lateral root initiation. They performed microarray transcript profiling on these roots and found that 4 stages precede pericycle cell division: G_1 cell cycle block, signal transduction, and progression through the G_1/S and G_2/M transitions. One of the genes observed in this study, *Krp2*, a cyclin-dependent kinase inhibitor that negatively regulates cell cycle activity and is transciptionally down regulated by auxin, was strongly expressed in most pericycle cells but repressed in cells in the early stages of lateral root initiation, indicating that KRP2 is involved in blocking the G_1/S transition in the pericycle cells not forming lateral roots (Himanen et al. 2002 and 2004). Other gene products involved in regulating the auxin signal in lateral root formation are NAC1, a transcriptional activator that is induced by auxin and promotes lateral root development (Xie et al. 2000), and SINAT5, a ubiquitin ligase that targets NAC1 for proteolysis, attenuating the auxin signal in pericycle cells not destined to become lateral roots (Xie et al. 2002). In addition, there are myriad other proteins and transcription factors involved in the complex signaling pathways in lateral root development.

Modes of Action of Plant Growth-Promoting Bacteria

Secretion of plant hormones

Many plant growth-promoting bacteria have been shown to secrete plant hormones, such as auxin, gibberellins, and cytokinins (Table 5-1). These hormones have been implicated as a

cause of plant growth promotion by bacteria since the effects of bacterial inoculation on root architecture (such as increased root hair and lateral root number) are similar to the effects of exogenously applying a solution of these hormones (Tien et al. 1979). Most research on plant growth promotion by bacteria resulting from the secretion of plant hormones has focused on auxin. It has been estimated that production of the auxin, indole-3-acetic acid (IAA), is common among bacteria in the rhizosphere, with approximately 80% being capable of synthesizing this hormone (Patten and Glick 1996).

Although there is a lot of corollary evidence for auxin's role in bacterial plant growth promotion, such as the similar effects bacterial inoculation and application of IAA have on plant roots, it does not exclude the possibility that the bacteria may be producing other growth-promoting factors that act similarly to or activate the same plant developmental pathways as auxin. In order to confirm that auxin is responsible for changes in root morphology, genetic analysis is also required. For example, when investigating how *Pseudomonas thivervalensis* causes shorter, more branched root systems, mutant analysis involving the screening of *Arabidopsis* mutants showed that plants insensitive to auxin did not experience the same root system shortening as wild type plants, indicating that auxin plays a role in this response (Persello-Cartieaux et al. 2001). Alternatively, other genetic analyses of auxin's role in bacterial growth promotion have focused on the bacterial side, looking at mutants with reduced auxin production. Most plant growth-promoting bacteria secrete auxin via the indole-3-pyruvate pathway, and the key enzyme in this pathway, indole-3-pyruvate decarboxylase, is the rate-limiting step (Lambrecht et al. 2000; Koga et al. 1994). This enzyme, encoded by the *ipdC* gene, is present in both Kp342 and another plant growth-promoting endophytic bacteria, *Enterobacter cloacae* P101 (Table 5-2). This gene has been the target for knocking out auxin

synthesis in several plant growth-promoting bacteria. In the case of *A. brasilense*, knocking out the *ipdC* gene reduces auxin production by 90%, resulting in a reduction of lateral root length and number, root hair formation, and primary root length inhibition compared to wild type strains (Barbieri and Galli 1993; Dobbelaere et al. 1999). Even in *ipdC* mutant experiments, the mutants still result in some changes in root architecture compared to uninoculated plants. This observation is likely due to the fact that the *ipdC* mutation does not completely abolish auxin production since there are multiple pathways for auxin synthesis that bacteria could use. For instance, an *ipdC* mutant of *Pseudomonas putida* still increased the number of adventitious roots formed by mung bean cuttings compared to uninoculated controls (Patten and Glick 2002). In addition, the effect of the *ipdC* mutation on auxin production in *A. brasilense* depends on the carbon source the bacteria are grown on, with mutants grown on pyruvate, lactate, or fumarate not producing significantly different amounts of auxin than wild type strains, suggesting an alternate tryptophan-dependent pathway of auxin synthesis (Carreno-Lopez et al. 2000). Since there are multiple pathways for auxin synthesis (Bartel 1997), it has proven difficult to obtain bacterial mutants in which auxin production has been completely abolished. As a result, the involvement of other yet uncharacterized growth-promoting signals secreted by the bacteria cannot be discounted.

As mentioned previously, plant growth-promoting bacteria have also been shown to produce gibberellins and cytokinins. In the case of gibberellin, Fulchieri and associates (1993) found that putting GA_3 on roots in concentrations similar to those produced by bacteria increased root growth in maize seedlings and that inoculation with different *Azospirillum* strains increased GA_3 levels in maize roots. Application of GA_3 to culture media also results in increased lateral root numbers in pearl millet plants (Tien et al. 1979). Although these results are intriguing, GA

synthesis genes were not identified or knocked out in the *Azospirillum* strain in order to directly demonstrate a role for gibberellin synthesis by the bacteria in the observed growth response. In the case of cytokinin, which is commonly believed to have an inhibitory effect on lateral root formation, kinetin and trans-zeatin inhibit the formation of new lateral root primordia, but increase the elongation of existing lateral roots in rice (Debi et al. 2005). If plant growth-promoting bacteria were producing this hormone, the elongation of existing root primordia may appear to be an increase in lateral root number when viewed by the naked-eye.

In addition to secreting plant hormones, some plant growth-promoting bacteria are known to produce enzymes that interfere with the formation of plant hormones. For example, ACC deaminase, an enzyme that lowers the concentration of ethylene in plants by cleaving its precursor, 1-aminocyclopropane-1-carboxylate (ACC), is found in several plant growth-promoting bacteria (Glick 2005). Since ethylene is known to inhibit root elongation (Abeles et al. 1992), bacteria that lower the levels of this hormone in the plant would be expected to result in plants with longer root systems. This role of ACC deaminase in altering root morphology has been supported by analysis of bacterial mutants. Plants inoculated with ACC deaminase producing strains of *P. putida* have longer roots than uninoculated plants while mutant bacteria lacking ACC deaminase do not increase root length (Hall et al. 1996). In addition, transferring the ACC deaminase gene into a *P. fluorescens* strain lacking this enzyme confers the ability to increase root length in canola while the wild type strain has no effect, suggesting that bacterial ACC deaminase lowers ethylene levels inside the plant, resulting in an increase in root length (Wang et al. 2000). Neither Kp342 nor *E. cloacae* P101 possess this enzyme (Table 5-2), so expressing this gene in these strains may further enhance their growth-promoting ability.

Production of volatile compounds

Some plant growth-promoting bacteria elicit increased growth in *Arabidopsis* by emitting volatile compounds (Ryu et al. 2003). *Bacillus subtilis* GB03 and *Bacillus amyloliquefacians* IN937a increase total leaf surface area in *Arabidopsis* when grown together on divided Petri dishes that allow only volatile signals to pass between the plants and bacteria (Ryu et al. 2003). Gas chromatography-mass spectrometry (GC-MS) analysis of the volatile compounds produced by these bacteria identified two compounds, 3-hydroxy-2-butanone (acetoin) and 2,3-butanediol, that appear to elicit the growth response (Ryu et al. 2003). Exogenous application of 2,3-butanediol onto plants and inoculation with bacteria blocked in the synthesis of these two substances confirmed that they play a role in plant growth promotion (Ryu et al. 2003). *E. cloacae* P101 and *E. coli* K12 both possess an operon responsible for the production of these two compounds, but Kp342 appears to have only one gene from this operon (Table 5-2). Since acetoin and 2,3-butanediol are products of the butanediol fermentation pathway present in many enteric bacteria (Madigan and Martinko 2006), these compounds may not be unique to plant growth promotion.

Making mineral nutrients more available to plants

Phytases from plant growth-promoting *Bacillus* strains degrade myo-inositol hexakisphosphate under low phosphate conditions, making phosphate available to plant roots from soil phytate, and Bacilli lacking a functional phytase do not elicit the same growth response as wild type strains (Idriss et al. 2002). Neither Kp342 nor *E. cloacae* P101 possess this enzyme (Table 5-2), though it could be a candidate for being engineered into these strains to further enhance plant growth promotion. Phosphate availability alters root architecture, with low phosphate conditions yielding higher numbers and densities of lateral roots (Lopez-Bucio et al. 2002; Williamson et al. 2001). In contrast to these results, Chevalier and associates (2003) found

that phosphorus starvation reduced primary root length and lateral root number. Although there is still ambiguity and contradiction over the effect of phosphate on root architecture, these observations indicate the possibility that plant growth-promoting Bacilli may be altering root architecture by increasing the amount of phosphate available to plants.

In addition to phosphate, nitrogen has also been shown to alter root architecture. Since many plant growth-promoting bacteria, including Kp342, are known to fix nitrogen, it is possible that nitrogen fixation may be involved in the modification of plant root systems by these bacteria. High sucrose to nitrogen ratios inhibit lateral root initiation (Malamy and Ryan 2001). In addition, localized low nitrate concentrations have been shown to increase lateral root formation while higher concentrations increase the rate of lateral root elongation (Zhang et al. 1999). Therefore, growth-promoting bacterial strains could also enhance growth by making nutrients more available to their plant hosts.

Secretion of other compounds

In addition to plant hormones and nutrients, plant growth-promoting bacteria may also be secreting other novel compounds that could be eliciting increased lateral root numbers and growth responses. For instance, the antibiotic 2,4-diacetylphloroglucinol (DAPG) produced by a strain of *Pseudomonas fluorescens* has been shown to increase root length and weight and "transiently" increase lateral root formation in pea plants, indicating this molecule may act similarly to plant hormones (De Leij et al. 2002). Another example of a compound that acts similarly to a plant hormone is affinin, an alkamide that occurs naturally in plants (Ramirez-Chavez et al. 2004). The exogenous application of this compound results in increased lateral root density and enhanced primary root and root hair growth in low concentrations (Ramirez-Chavez et al. 2004). Although these effects are similar to those of auxin, experiments using auxin-inducible genes fused to a β-glucuronidase (GUS) reporter and auxin-resistant

mutants have demonstrated that affinin's mode of action is independent of auxin (Ramirez-Chavez et al. 2004). Amidenin, another alkamide that is produced by the fungus *Amycolatopsis* sp., has been shown to promote growth in rice seedlings (Kanbe et al. 1993). Therefore, the presence of other plant growth-promoting compounds in bacteria cannot be discounted, even when the plant growth response is similar to that seen in response to plant hormones like auxin.

Results

Increased Lateral Root Number in *Arabidopsis thaliana*

Kp342 was first observed to increase lateral root number during an experiment to determine its route of entry into plants, which was hypothesized to be through cracks formed at the base of lateral roots. To test this hypothesis, Kp342 was inoculated onto an *Arabidopsis* mutant, *xbat-32*, which has a reduced number of lateral roots (Nodzon et al. 2004), to see if the bacteria colonized the mutant to a lesser extent than wild type plants. However, after inoculation, the number of lateral roots on inoculated *xbat-32* mutants was increased relative to uninoculated controls, nearly rescuing the mutant phenotype (Figure 5-1b). In addition, the wild type plants also displayed an increase in lateral root number (Figure 5-1a, b). Once observed, this lateral root increase by Kp342 indicated a possible cause of this bacterium's nitrogen fixation independent growth promotion.

Lateral Root Increase Not Due to Nitrogen Fixation

Before making any attempts to identify the lateral root-increasing signal produced by the bacteria, it was necessary to first confirm that this growth response was not a result of nitrogen fixation, since changes in nitrate concentrations are known to alter lateral root development. To see if nitrogen fixation by Kp342 was responsible for its lateral root increasing phenotype, a *nifH* mutant of the bacteria which does not fix nitrogen in plant tissue (Iniguez et al. 2004) was

inoculated onto *Arabidopsis* and found to increase lateral root numbers compared to uninoculated plants (Figure 5-2). Although the plants inoculated with the *nifH* mutant showed a slightly reduced number of lateral roots compared to those inoculated with wild type Kp342, this difference was small and possibly due to enhanced nutrition provided to the plants in the form of fixed nitrogen produced by Kp342. As a result, it was concluded that the primary cause of the lateral root increase was not nitrogen fixation.

Lateral Root Increasing Phenotype is Strain Specific

In addition to determining if the lateral root increase was not due to of nitrogen fixation, it was also important to determine if this trait is common among enteric bacteria, or specific to Kp342 and other plant growth-promoting bacteria. In order to determine this, Kp342 was compared with other enteric bacteria, including *E. coli* K12 and *K. pneumoniae* type strain 13883 (Kp13883) obtained from the American Type Culture Collection (ATCC). *E. coli* K12 was used to test if other enteric bacteria possess the genes required for eliciting a lateral root increase and the type strain of *K. pneumoniae* was used to test if this trait is common to the *Klebsiella* genus. The results of these inoculations demonstrated Kp342 increased lateral root number more than Kp13883, while *E. coli* inoculated plants were not statistically different from uninoculated plants (Figure 5-3a). Although, Kp13883 slightly increased lateral root numbers to a certain extent, it was significantly less than Kp342.

In addition, Kp342 inoculation was also compared to inoculation with another plant growth-promoting bacteria, *E. cloacae* P101 (Riggs et al. 2001), in order to determine if this trait is present in other enteric plant growth promoters. The results from this experiment demonstrated that P101 does indeed increase lateral root number to the same extent as Kp342 (Figure 5-3b). These results indicate that the lateral root increasing phenotype is not a ubiquitous trait among other enteric bacteria, but is more commonly found in strains isolated from plant

tissue, possibly developed by the bacteria as an adaptation to living in association with plants. Although the type strain, Kp13883, slightly increased lateral root number, the intensity of this response is specific to plant growth-promoting strains like Kp342 and P101. The difference in the intensity of the response may be due to different regulation of growth-promoting genes in these bacteria or multiple plant growth-promoting pathways present in Kp342 and P101 that are absent in Kp13883.

Lateral Root Promotion Due to Secreted Diffusible Product

Another basic question about the growth-promoting phenotype was whether Kp342 elicits the lateral root increase as a response to a secreted and diffusible compound produced by the plant or whether the bacteria must be present within plant tissue to elicit a response. To test if the bacteria were secreting a diffusible product responsible for increasing lateral root number, Kp342 was inoculated onto plant roots covered by a 0.2 μm membrane. This membrane served to prevent the bacteria from coming into direct contact with the roots while still allowing chemical signals to pass between the plants and the bacteria (Figure 5-4a). As controls, plants were also treated with only the membrane or with Kp342 inoculated under the membrane. Since Kp342 increased lateral root number to the same extent both under and on top of the membrane, it was concluded that it does secrete a product that can diffuse through the membrane to increase lateral root formation on the plants (Figure 5-4b). The results from this experiment also showed that the membrane itself altered lateral root development. Uninoculated plants under the membrane had higher numbers of lateral roots than uninoculated plants without a membrane (Figure 5-4b). In addition, inoculated membrane-covered plants had lower lateral root numbers than plants treated with Kp342 alone (Figure 5-4b). The effects the membrane had on root numbers were probably due to the fact that the membrane limited the amount of light and the gas exchange the root systems were exposed to.

Secretion of Lateral Root Promoting Compound is Plant Inducible

In the first attempt to isolate growth-promoting compounds secreted by Kp342, concentrated extracts from culture supernatants of overnight broths were assayed for lateral root increasing activity. Initially, the supernatant from Kp342 grown on Luria-Bertani (LB) broth was filter sterilized and fractionated on C18 resin with increasing concentrations of methanol to isolate lateral root promoting compounds. These fractions were then concentrated on a rotary evaporator so that the concentration of any growth-promoting compounds would be high enough to elicit a growth response, since previous experiments showed unconcentrated Kp342 culture supernatants do not elicit a growth response. These LB supernatant fractions were assayed on plants to check for lateral root induction compared to Kp342 inoculation. All the fractions appeared to increase lateral root numbers to some degree but a subsequent experiment yielded conflicting and variable results, demonstrating that more controlled bacterial culture conditions were required to reliably elicit the production of lateral root inducing compounds. In order to analyze a more controlled growth environment, another experiment was performed with murashige and skoog (MS) media, a defined plant growth media that Kp342 grows on after inoculation. Use of this defined media simplified the composition of the culture extract for easier analysis. When plants were treated with the concentrated MS culture supernatant, they did not increase lateral root numbers compared to plants treated with concentrated media (Figure 5-5). Since MS is a defined media, the possibility that it may not contain the substrate required to produce growth-promoting compounds was considered. The bacteria may require certain substrates or signals from the host plant in order to produce compounds that elicit a plant growth response.

Lateral Root Promotion May Involve Secretion of a Plant Hormone

When lateral root promotion was first observed in response to Kp342 inoculation, the immediate candidate suspected to be involved was auxin. The Kp342 genome has been sequenced and the annotation has identified the presence of an *ipdC* auxin synthesis gene. The observations that Kp342 possesses this auxin synthesis gene and that another strain of *K. pneumoniae* has been shown to secrete auxin (El-Khawas and Adachi 1999), coupled with the fact that exogenous application of auxin is known to increase lateral roots on plants, further substantiates the hypothesis that Kp342 may be producing auxin to elicit this lateral root promotion. In an attempt to confirm if auxin and other plant hormones are involved in this response, *Arabidopsis* mutants described as being insensitive to certain plant hormones were tested to see if any of these mutations abolished the plant's response to Kp342 (Table 5-3). When assayed, most of these mutants responded to Kp342 inoculation (Figure 5-6). Unfortunately, the response of these mutants to exogenous application of the hormone they are insensitive to has not been studied in regards to lateral root formation. Most were only examined for reduced inhibition of primary root length. As a result these data must be interpreted with caution, since these genes may not be involved to lateral root formation. For example, the *axr2-1* mutant is insensitive to auxin in that its primary root elongates in the presence of high auxin concentrations, but it is sensitive to auxin's effect on lateral root formation, forming more lateral roots when treated with auxin (Knee and Hangarter 1996). Therefore, these hormones could not be discounted as being involved in lateral root promotion. In order to obtain more conclusive data, mutants insensitive to auxin that do not significantly increase lateral root number in response to exogenous hormone application were selected for further experiments. One of these mutants, *nph4-1*/*arf19-4* (Wilmoth et al. 2005), was tested and did not increase

lateral number after Kp342 inoculation, indicating that auxin may be involved in the lateral root response (Figure 5-6).

Exogenous Application of Auxin Partially Mimics Lateral Root Promotion by Kp342

With the difficulty in extracting growth-promoting compounds from Kp342 supernatants and the uncertainty of the conclusions drawn from lateral root promotion on auxin insensitive mutants, Kp342 inoculation was compared directly to exogenous application of auxin on wild type Col-0 *Arabidopsis* plants. The root systems of Kp342 inoculated plants were compared to their auxin treated counterparts in regards to both lateral root number and primary root length. When looking at the number of lateral roots produced by these plants, both Kp342 inoculation and treatment with 0.1 μM IAA increased lateral root number to the same extent (Figure 5-7a). In contrast, exogenously applied auxin significantly inhibited primary root length while Kp342 inoculated plants did not significantly differ from uninoculated negative controls in that regard (Figure 5-7b). Therefore, exogenous auxin application partially, but not completely, mimics the effects of Kp342 inoculation.

Role of *ipdC* in Auxin Production

Initially, attempts were made to delete the *ipdC* gene in Kp342. However, the construction of in-frame deletions in Kp342 proved to be problematic due to a variety of problems. These problems include the presence of a wide range of antibiotic resistance genes commonly used in gene deletion protocols and the inefficient transformation efficiency of the Kp342 strain. As a result, further genetic work on determining the role of the *ipdC* gene and IAA production on lateral root promotion was carried out in *E. cloacae* P101. As P101 also increases lateral root numbers in *Arabidopsis* similar to Kp342 (Figure 5-3b) and its genome annotation revealed the presence of the *ipdC* gene, it served as an ideal substitute.

An in-frame deletion of the *ipdC* gene in P101 was constructed using a modified Red recombinase-mediated homologous recombination method (Jantama et al. 2008). Briefly, the *ipdC* gene was first replaced with an antibiotic resistance cassette via a double-crossover homologous recombination, which was later removed by an additional double-crossover recombination event with a deoxyribonucleic acid (DNA) fragment containing a clean, in-frame deletion of the gene. When the culture supernatant of the Δ*ipdC* mutant was assayed, IAA production was lowered to 1 μg/mL, nearly 4-fold lower than the wild type P101 strain (Figure 5-8). This loss of auxin production in the Δ*ipdC* mutant supported its proposed role in auxin production by P101. To further confirm the role of *ipdC* in IAA production, the Δ*ipdC* mutant was complemented with a plasmid-born copy of *ipdC* on pHLT14. When assayed, IAA production in the complemented Δ*ipdC* strain was restored, with a concentration eight times higher than the wild type strain (Figure 5-8). The fact that the complemented mutant produced eight times more auxin than wild type P101 is probably due to the higher copy number of the plasmid relative to the single copy number of the *ipdC* gene found on the chromosome. Regardless, restoration of auxin production in Δ*ipdC* by pHLT14 further confirmed the *ipdC* gene's role in auxin production in P101.

pHLT14 was constructed to include the *par* locus for plasmid stability so that it could be maintained by bacteria inside plant tissue in the absence of selection pressure. To verify stability of the plasmid in the absence of selection pressure, the numbers of antibiotic resistant and sensitive cells were counted after two subculturings in LB media without antibiotics. pHLT14 cultures had the same number of antibiotic resistant and sensitive cells, indicating the bacteria had retained the plasmid (Figure 5-9). In comparison, cultures carrying a complement plasmid lacking the *par* locus had significantly fewer antibiotic resistant cells than non-resistant cells,

indicating the plasmid was lost (Figure 5-9). Therefore, the *par* locus served its intended purpose of maintaining pHLT14 in Δ*ipdC* in the absence of selection pressure.

Δ*ipdC* Mutants on Plants

When the Δ*ipdC* mutant in P101 was included in experiments on *Arabidopsis* plants, no differences were seen in lateral root numbers between any treatment, even uninoculated and P101 wild type treated plants. With the loss of this phenotype, the role of *ipdC* catalyzed auxin production in lateral root promotion previously observed in response to bacterial inoculation could not be verified. In the event that loss of the phenotype was due to a change in P101, Kp342 was retested on plants since the possibility of two bacterial strains losing the same phenotype was unlikely. Neither wild type P101 nor Kp342 increased lateral root numbers when inoculated on *Arabidopsis* Col-0 as previously observed, indicating loss of lateral root induction was not due to a change in the bacterium and, instead, might be due to an artifact in plant culturing. Since the loss of the lateral root inducing phenotype was roughly preceded by the purchase of new MS media, the possibility that the media had been contaminated was considered; therefore, new MS media was procured. Even when new media was used, plants still did not increase lateral root numbers in response to bacterial inoculation relative to uninoculated plants. Because the lateral root response was first observed in *xbat-32* mutants and the intensity of the response was more dramatic in this mutant, experiments were repeated on this *Arabidopsis* line, but again, inoculated and uninoculated plants did not differ from each other. In all new cases, uninoculated plants were seen to have more lateral roots than previously observed. Due to the inability to restore lateral root promotion by changing the strain, plant culture media, or *Arabidopsis* line, it is believed the loss of this phenotype is due to gradual changes and improvement in handling of the plants over time.

Discussion

The endophytic bacterium, *K. pneumoniae* 342, is of interest as a model endophyte due to its ability to colonize the interior of plant tissues and increase plant growth yield. While this endophyte has been found to provide fixed nitrogen to plants and results in a generalized enhancement in plant growth, further studies on Kp342 found that it could enhance lateral root growth in *A. thaliana* by a mechanism independent of nitrogen fixation. This lateral root-enhancing phenotype was only observed in other plant growth-promoting bacteria and deemed to be an ideal assay for identifying the mechanism of Kp342's nitrogen fixation independent growth-promoting mechanism. As outlined earlier in this chapter, plant growth-promoting bacteria possess a variety of mechanisms that enhance plant growth. Attempts to isolate growth-promoting compounds from Kp342 culture supernatants failed, so the genome sequences of Kp342 and another endophytic bacterium, *E. cloacae* P101, were examined for genes associated with these mechanisms. As a result of this examination, auxin production by indole-pyruvate decarboxylase was deemed to be an ideal candidate for further study of these two endophyte's growth-promoting effect on plants.

Initial attempts to determine if auxin or another plant hormone plays a role in Kp342's lateral root promotion involved assaying a wide array of hormone insensitive mutants for their response to bacterial inoculation. It was expected that if Kp342 were eliciting a growth response by secreting one of these hormones, that response would be diminished or eliminated in a mutant insensitive the hormone. In nearly every mutant tested, Kp342 increased the number of lateral roots on plants. While this observation initially indicated that the growth-promoting response was due to a factor other than one of the hormones examined, the complexity of hormone signaling in *Arabidopsis* confounds the interpretation of these results. Phytohormones play multiple roles in plant growth and development. As a result, there are many different receptors

that respond to phytohormones in different signaling pathways, leading to different effects on plant biology. Therefore, a mutant that is insensitive to one effect of a hormone may still respond to that hormone in other developmental pathways. This fact was demonstrated by Knee and Hangarter (1996), who found several supposedly auxin insensitive *Arabidopsis* mutants that still increased lateral root formation when the hormone was applied to plant growth media. As most of the hormone insensitive mutants used in this study were originally assayed for insensitivity based on root elongation rather than lateral root response to hormone application, the increase in lateral root numbers seen on these mutants is inconclusive. Therefore, mutants specifically insensitive to hormone application in regards to lateral root production were selected. The search for such mutants proved to be difficult owing to the essential and redundant nature of genes involved in lateral root initiation (Malamy and Ryan 2001). One of the few lateral root specific mutants obtained was *nph4-1/arf19-4* (Wilmoth et al. 2005). This auxin insensitive mutant did not increase lateral root numbers in response to Kp342, indicating that auxin production by Kp342 could be the cause of lateral root promotion. It should be noted that, given the severity of lateral root initiation mutants, it is possible that no growth-promoting compound could rescue it. Therefore, more evidence was required to confirm the role of auxin production in Kp342-mediated lateral root promotion.

Interestingly, while Kp342 increased lateral root number similar to exogenous auxin application, it did not inhibit primary root length as was observed upon auxin treatment. One possible reason for the partial mimicry of auxin application is that Kp342 secrets auxin at such a low level that it promotes lateral root branching, but does not inhibit root elongation. Because Kp342 is present in and on the plants, it could produce a constant source of auxin so that even low levels produced by the bacteria could increase lateral root growth without reaching an

inhibitory concentration. Exogenous application of IAA involves treating plants with higher concentrations in the media, since low concentrations may breakdown or be metabolized by the plant over time. As a result, exogenous IAA could not completely simulate auxin production by Kp342, so differences between bacterial inoculation and exogenous auxin application are not unexpected. Therefore, auxin remains a likely candidate for Kp342's lateral root promoting effect on plants.

In order to confirm auxin's role in lateral root promotion, experiments involving strains abolished in auxin production were needed. Analysis of the Kp342 genome indicated the presence of the *ipdC* auxin synthesis gene, which was selected as the target for knocking out auxin production in this strain. Unfortunately, genetic manipulations in Kp342 proved to be problematic. The strain possesses resistances to multiple antibiotics commonly used as selectable markers in molecular biology, including ampicillin, tetracycline, and chloramphenicol (Appendix A), so special care had to be taken in selecting plasmid vectors to use in gene knockout approaches. The approach selected to delete the *ipdC* gene is a Red recombinase-mediated homologous recombination method (Jantama et al. 2008). Unfortunately, Kp342 was unable to be transformed with a plasmid carrying the Red recombinase genes. This problem was initially believed to be due to plasmid incompatibly with one of Kp342's native plasmids, but several attempts to transform the bacteria with the Red recombinase genes on other plasmids with different origins of replication also failed. Further investigation determined that the problem lies with the arabinose induction and Red recombinase genes themselves. The entire set of Red recombinase genes plus the upstream arabinose induction genes were cloned into pCR4.1, a vector previously transformed into Kp342, but transformation of the recombinase carrying pCR4 still yielded no transformants. Therefore, it appears that either the Red

recombinase or arabinose induction genes themselves are incompatible with Kp342. In addition to this difficulty, Kp342 also has a low transformation efficiency, which is a common trait in members of the *Klebsiella* genus (Regue et al. 1992; Fournet-Fayard et al. 1995). To ensure that the screening of transformants for mutants in any gene deletion method was successful, high transformation efficiencies would be needed. Therefore, further attempts to make gene deletions in Kp342 were halted in favor of working with a more genetically tractable bacterium.

Because P101 enhances lateral root growth similar to Kp342 and also possessed the *ipdC* gene, it was selected for genetic studies on the role of auxin production in the lateral root response. P101 proved to be more amenable genetic manipulation than Kp342 and an Δ*ipdC* mutant in this strain was constructed with relative ease. This mutant was deficient in auxin production compared to the wild type strain, and that production was restored upon complementation with a plasmid-born copy of the gene. Unfortunately, once the Δ*ipdC* mutant was constructed, the lateral root growth response seen on plants had diminished and significant differences in lateral root numbers between treatments were no longer observed, not even between uninoculated and wild type P101 or Kp342 treated plants. The sudden loss of this previously reliable and reproducible lateral root promoting phenotype was baffling, and repeated troubleshooting measures were unable to identify the source of the problem. Because the phenotype was no longer observed in response to either P101 or Kp342, loss of lateral root induction due to evolution of the bacteria was discounted, as it is unlikely that two different strains would lose the same phenotype. Therefore, problems stemming from plant culturing were considered. The loss of the lateral root induction roughly followed the purchase of new MS media, so the possibility that the media had been contaminated by a synthetic auxin was considered. If auxin were present in the media, even uninoculated plants would be stimulated to

produce lateral roots, confounding any results comparing treated and untreated plants. But even on new media, lateral root induction was still not observed. One final possibility for the loss of lateral root promotion was handling of the plants and human error. Working with *Arabidopsis* can be difficult owing to the small and delicate nature of the plant's root system, but over time handling of the plants became easier with practice and experience. It is possible that plants in earlier experiments were stressed by handling, and such stress could have altered root development in the plant. After handling of the plants became more routine, it is possible that they were less stressed and root development less responsive to bacterial inoculation. Therefore, the changes in the lateral root numbers produced by plants may be an artifact of gradual changes in handling that occurred over time.

Regardless of the loss of the lateral root promoting phenotype, the work in this chapter also resulted in the construction of plasmids carrying known plant growth-promoting genes that could have potential uses in future studies. The *ipdC* gene, known to be involved in growth promotion by several plant-associated strains, was cloned into a low copy number plasmid that was modified with a plasmid stability locus, allowing it to be maintained by bacteria in plant tissues in the absence of selection pressure. In addition, though not utilized in the current experiments, another plasmid carrying the *acdS* gene encoding ACC deaminase, a known plant growth-promoting factor, was also constructed. This *acdS* gene was placed immediately downstream of the *ipdC* gene in the original complement plasmid so that both genes would be under control of the same promoter. This new plasmid construct with two known plant growth-promoting genes could be transformed into other plant-associated bacteria and potentially increase positive effects on plant growth. In addition, since ACC deaminase acts by breaking down the precursor of the plant hormone ethylene, which is known to act in plant

defense pathways (Glick 2005; Iniguez et al. 2005; Pieterse et al. 1998), this plasmid may also lower plant defense responses to endophytic bacteria, resulting in increased colonization and a concomitant increase in growth promotion. Therefore, the plasmids constructed in this study have potential for future studies in plant growth-promoting bacteria.

Materials and Methods

Plant Culturing

A. thaliana seeds were obtained from the Arabidopsis Biological Resource Center, with the exception of *nph4-1/arf9-4*, which was provided by Jason Reed at the University of North Carolina, and *xbat-32*, which was provided by Wen-Yuan Sun at the University of Florida. Arabidopsis seeds were surface sterilized by submersion in 70% ethanol for 5 minutes (min), rinsing 4 times in sterile water, submersion in 10% bleach for 20 min, and rinsing 4 more times in sterile water. Seeds were germinated on half strength MS media (Sigma) with 10 g L^{-1} sucrose and 0.9% noble agar and cold treated at 4^{o}C in the dark for 3 days. After cold treatment, seeds were transferred to a 22^{o}C growth chamber with an 11 hour daylight cycle. Plates with seeds were oriented vertically in the chamber so plant roots would grow along the surface of the agar. After 8 days, seedlings were transferred to MS media without sucrose and treated 1 day later. Lateral root number and primary root length were counted 9 days after inoculation. All experiments assaying lateral root growth used 30 plants per treatment.

Bacterial Strains and Inoculum Preparation

Bacterial strains used were *K. pneumoniae* 342, *K. pneumoniae* ATCC13883, *E. cloacae* P101, and *E. coli* K12. All bacteria were cultured on Luria-Bertani (LB) medium at 37^{o}C, with the exception of strains carrying pLOI341, which were cultured at 30^{o}C. Inocula were prepared by scraping cells from LB agar plates and suspending them in sterile phosphate buffered saline (PBS). Bacteria were then diluted to 10^4 colony forming units (CFU)/mL and 10 μL were

inoculated per plant. To test for a secreted product, a sterile membrane filter with a 0.2 μm pore size was placed over the root system of plants and bacteria were inoculated on top. As a positive control, bacterial inoculum was applied underneath the membrane. Uninoculated plants with just the membrane were used as a negative control.

Extraction of Compounds from Culture Supernatant

Kp342 was grown in 4 L of LB broth or MS broth overnight at 28^{o}C with shaking. Culture supernatants were collected after centrifuging at 5000 rpm for 30 min. Additional bacteria were removed from the supernatant by filtering through a 0.2 μm membrane. The LB supernatant was fractionated on C18 resin with increasing concentrations of methanol. Briefly, the LB supernatant was poured over C18 and solutes were eluted with 500 mL each of 25, 50, 75, and 100% methanol. The MS supernatant and LB methanol fractions were concentrated on a Buchi Rotavapor at 90 rpm, with the water bath set at 50^{o}C and the condenser cooled with water chilled to 4^{o}C. The volume of each sample was concentrated down to 5 mL and stored at -80^{o}C until use in plant assays.

Comparison of Bacterial Inoculation to Kp342 Supernatant Extracts and Exogenous IAA

Col-0 *Arabidopsis* plants were cultivated and inoculated as described previously. For comparison to concentrated culture supernatants, plants were treated with 10 μL of supernatant extracts instead of bacterial suspensions. For comparison to exogenous application of auxin, varying amounts of IAA were added to MS media from a 1 mM IAA stock solution in 100% ethanol immediately prior to pouring plates. The resulting IAA concentrations in the media were 0.1, 1, and 10 μM IAA. MS agar with plain ethanol added in the same volume as the IAA stocks was used as a negative media control. Plants were transferred to fresh media 9 days after germination, and were either inoculated with Kp342 as described previously or transferred to

plates with IAA containing media. Lateral root number and primary root length were measured 9 days after inoculation.

Construction of *ipdC* In-Frame Deletion in *E. cloacae* P101

P101 DNA was extracted using the FastDNA Kit. All plasmids were extracted using a QIAprep Spin miniprep kit. The *ipdC* gene in *E. cloacae* P101 was knocked out using a method involving two double-crossover homologous recombinations with linear, polymerase chain reaction (PCR) -generated constructs (Jantama et al. 2008). The *ipdC* gene plus 300 basepairs (bp) up and downstream was PCR amplified with Ex Taq (Takara) using primers A_P101_ipdC and D_P101_ipdC (Table 5-4). Reactions contained 5 μL of 10x Ex Taq buffer, 200 μM of each dNTP, 500 ng of P101 DNA, 2 μM of each primer, and 1.25 units of Ex Taq in 50 μL. Cycle conditions were as follows: 94^{o}C for 30 seconds (s), 30 cycles of 94^{o}C for 10 s, 55^{o}C for 30 s, and 72^{o}C for 2 min 40 s, followed by a 5 min final extension at 72^{o}C. The resulting PCR product was cloned into the pCR4.1 TOPO cloning vector, yielding pHLT4. An inverted PCR of the pHLT4 plasmid was performed with Phusion polymerase (New England Biolabs) using primers F2-JMP101ipdC and R2-JMP101ipdC (Table 5-4). The reaction contained 10 μL of 5x HF buffer, 200 μM of each dNTP, 500 ng of pHLT4, 2 μM of each primer, and 2 units of Phusion polymerase in a final volume of 50 μL. The reaction was cycled under the following conditions: 98^{o}C for 30 s, 30 cycles of 98^{o}C for 10 s, 58^{o}C for 30 s, and 72^{o}C for 1 minute 15 s, followed by a 5 min final extension at 72^{o}C. The primers in this reaction were designed to include the start and stop codons of the *ipdC* gene and yield a product containing the plasmid flanked by the up and downstream regions of *ipdC*. This product was blunt ligated with a *cat/sacB* cassette excised from pLOI4162 (Jantama et al. 2008). The resulting plasmid, pHLT6, was used as a template for a PCR with primers A_P101_ipdC and D_P101_ipdC as described above, yielding a linear

fragment containing the up- and downstream *ipdC* regions with the *cat/sacB* cassette inserted in place of the gene. This linear construct was electroporated into P101 cells expressing Red recombinase from plasmid pLOI3421 (Wood et al. 2005). Red recombinase expression was under control of the arabinose promoter and induced in P101 by growth in media containing 5% arabinose. Insertions into *ipdC* were selected by screening transformants for chloramphenicol resistance and sucrose sensitivity. To remove the selectable marker, the *cat/sacB* cassette was digested out of pHLT6 using PacI. The plasmid was then religated, leaving a short scar sequence with stop codons in all 6 reading frames. This new plasmid, pHLT7, was used as template in another PCR with primers A_P101_ipdC and D_P101_ipdC to generate a new linear fragment with the up and downstream regions of *ipdC* flanking the stop codon scar sequence. This fragment was electroporated into P101 Δ*ipdC::cat/sacB* expressing Red recombinase. Resulting transformants were selected for sucrose resistance and screened for chloramphenicol sensitivity, insuring loss of the *cat/sacB* cassette in the genome and a non-polar deletion of the *ipdC* gene. The red recombinase carrying plasmid, pLOI3421, which has a temperature sensitive origin of replication, was cured from the strain by growth at 37^{o}C overnight, yielding *E. cloacae* P101 Δ*ipdC*.

Complementation of P101 Δ*ipdC*

The Δ*ipdC* mutant of P101 was complemented with a plasmid-born copy of the *ipdC* gene. The vector selected for complementation was pAYC184 (Chang and Cohen 1978). To construct the complement plasmid, the *ipdC* gene was amplified using primers A_P101_ipdC and D_P101_ipdC. These primers amplified 300 bp upstream of the gene to ensure the native promoter was obtained. Each primer contains a BamHI restriction site for easy cloning. The *ipdC* PCR product was digested with BamHI and then ligated into the BamHI site in the

tetracycline gene of pACYC184, yielding pHLT12 (Figure 5-10a). Transformants were screened for chloramphenicol resistance and tetracycline sensitivity to insure insertion of the *ipdC* gene into the plasmid.

In order for the plasmid to be maintained in plant tissue in the absence of selection pressure, the partitioning genes for plasmid stability from pTR102 were inserted (Weinstein et al. 1992). To insert the partitioning genes into pHLT12, the *par* locus was first amplified from pTR102 with Phusion polymerase using primers par1sa and par2sa (Table 5-4). The resulting PCR product was blunt ligated into the ScaI site in the ampicillin gene of pHLT4, yielding pHLT13. Primers tpkanpar1 and tpkanpar2 (Table 5-4) were then designed to amplify the region of pHLT13 that includes the inserted *par* locus and the adjacent kanamycin resistance gene. This step was performed to obtain a *par-kan* cassette that could be used as a selectable marker for the insertion of the *par* locus. The *par-kan* cassette from pHLT13 was amplified with Phusion polymerase and used as the insert in a blunt ligation into the ScaI site of pHLT12, yielding pHLT14 (Figure 5-10b). Transformants were screened for kanamycin resistance.

Determination of Plasmid Stability

Stability of pHLT14 was determined by growing Δ*ipdC* with pHLT14 in the absence selection pressure. Δ*ipdC* with pHLT12, which lacks the *par* locus for plasmid stability, was used as a negative control. Δ*ipdC* pHLT14 was streaked out on LB plates with 50 μg/mL kanamycin and Δ*ipdC* pHLT12 was streaked out on LB with 20 μg/mL chloramphenicol. One colony from each plate was used to inoculate 3 mL of plain LB. Cultures were grown at 37^{o}C with shaking for 24 hours and 30 μL were taken to subculture into new 3 mL broths. After 2 subculturings, dilution spread-plating was performed on plain LB and LB with the appropriate

antibiotic. Colonies on each media type were counted to determine the number of CFU/mL in each culture that were still resistant, and therefore, still carried the plasmid.

Auxin Assay

P101, Δ*ipdC*, and Δ*ipdC* pHLT14 cultures were grown at 30°C in M9 minimal media supplemented with 100 μg/mL tryptophan. At 12 hours intervals during the growth curve, culture OD_{620} was measured. Two milliliters of culture supernatant were collected by centrifugation at 13,000 rpm for 1 minute and filter sterilized with a 0.2 μm membrane to remove remaining cells. Supernatants were assayed for IAA using Salkowski's colorimetric assay (Gordon and Weber 1951). The Salkowski reagent was composed of 1 mL of 0.5 M $FeCl_3$ and 50 mL of 35% $HClO_4$. For the assay, 0.5 mL of Salkowski reagent was added to 1 mL of culture supernatant, vortexed, and incubated at room temperature for 20 minutes. Absorbance was measured at 530 nm on a Shimadzu UV-160 spectrophotometer. Concentration of IAA in culture supernatants was determined by comparison to a standard curve of 1, 2.5, 5, 7.5, 10, 15, 20, 30, 50, and 75 μg/mL IAA diluted in M9 media with 100 μg/mL of tryptophan.

Insertion of ACC Deaminase into pHLT14

ACC deaminase was added to pHLT14 by inserting the *acdS* gene from *P. putida* UW4 immediately upstream of *ipdC* so both genes would be under control of the same promoter. To do this, primers ACCd-FXhoI and ACCd-RSbfI (Table 5-4) were designed to amplify the coding region of *acdS*, including the ribosome binding site. Primers pACYCipdcRXhoI-1 and ACCd-RSbfI (Table 5-4) were designed to amplify pHLT14, excluding the upstream region of *ipdC* to eliminate any transcriptional terminators downstream of *ipdC*. In order for the *acdS* gene to insert in the same orientation as *ipdC*, the forward and reverse primers in each pair were designed to include an XhoI or SbfI restriction site, ensuring that the gene could only insert in

one orientation. The pHLT14 and *acdS* PCR products were digested with XhoI and SbfI and ligated together, resulting in plasmid pHLT16 (Figure 5-10c).

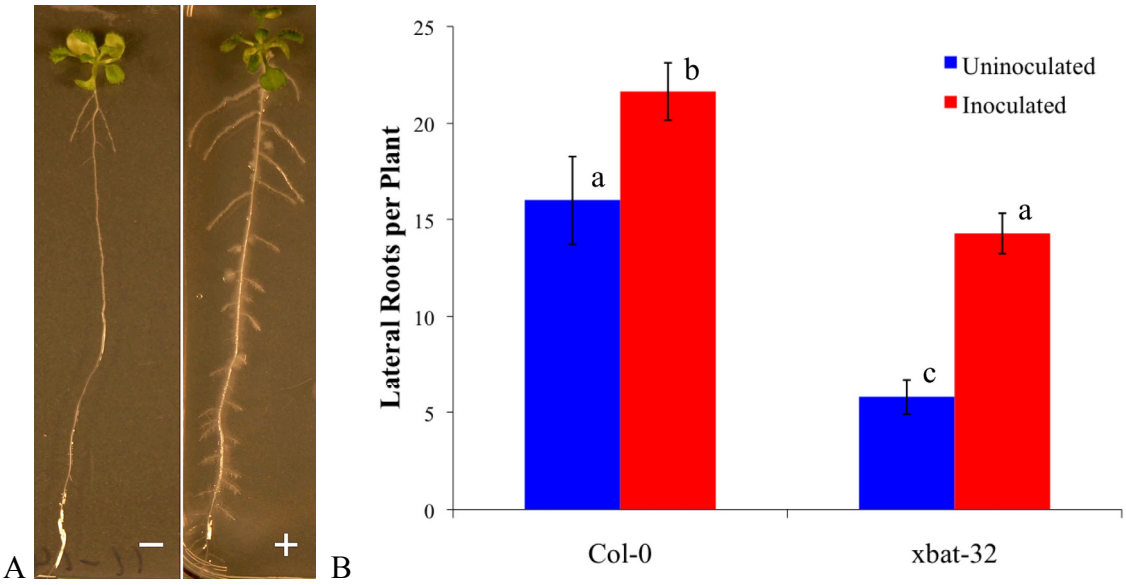


Figure 5-1. Lateral root promotion by Kp342. A) Uninoculated and Kp342 inoculated *Arabidopsis* Col-0. B) Lateral root numbers on Kp342 inoculated *Arabidopsis* compared to uninoculated plants in Col-0 and *xbat-32*. Plants were inoculated with 10 cells 9 days after germination and roots counted 9 days post inoculation. Statistical difference denoted by different letters.

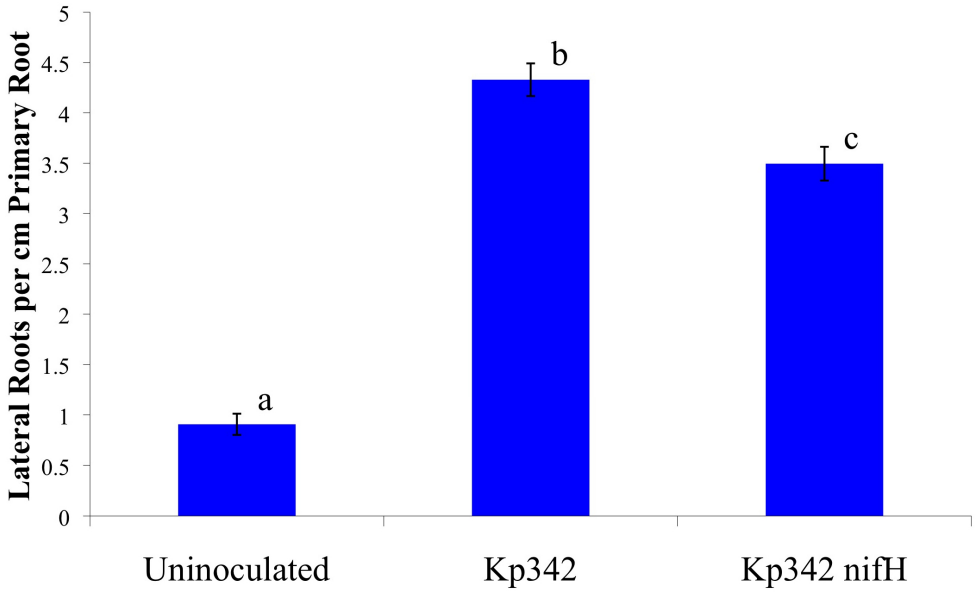
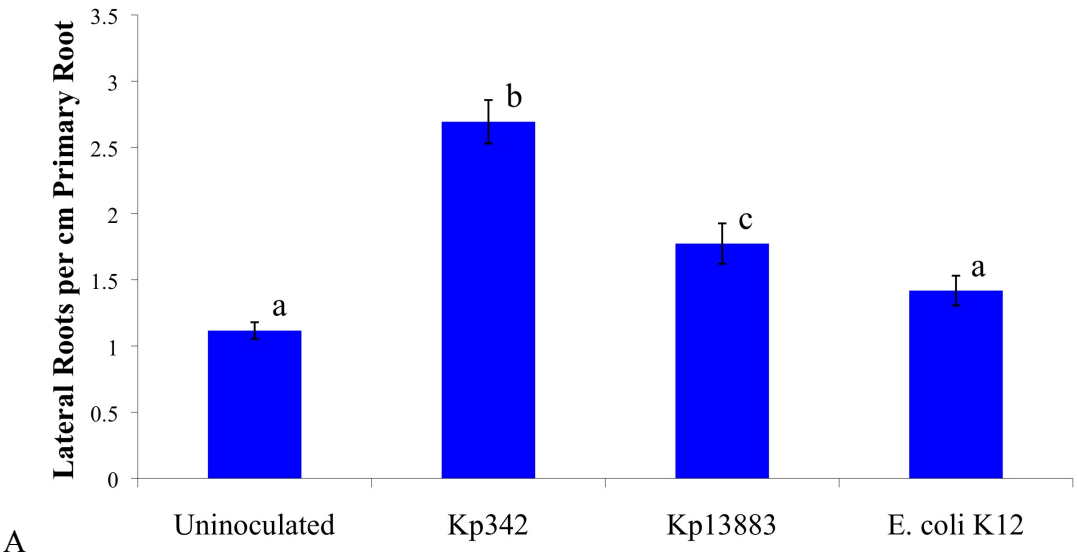
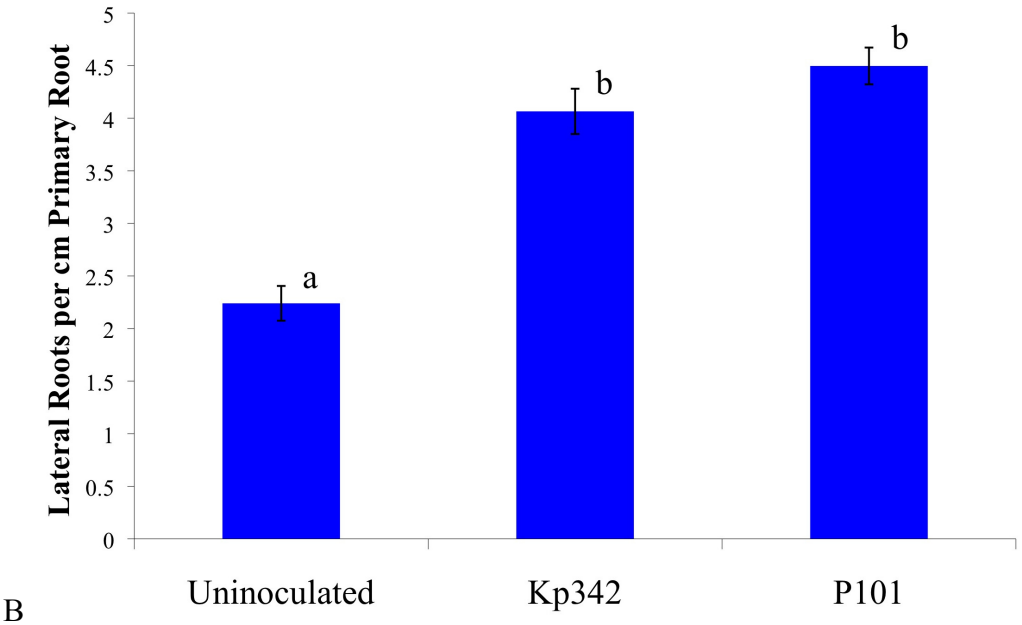


Figure 5-2. Effect of nitrogen fixation on lateral root promotion. Lateral roots per cm primary root on *Arabidopsis* inoculated with wild type Kp342 or a *nifH* mutant compared to uninoculated plants. Plants were inoculated with 10 cells 9 days after germination and roots counted 9 days post inoculation. Statistical difference denoted by different letters.

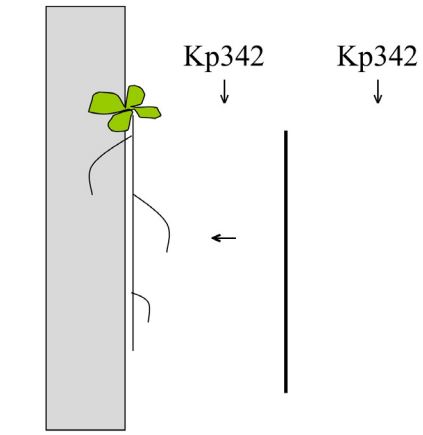


A

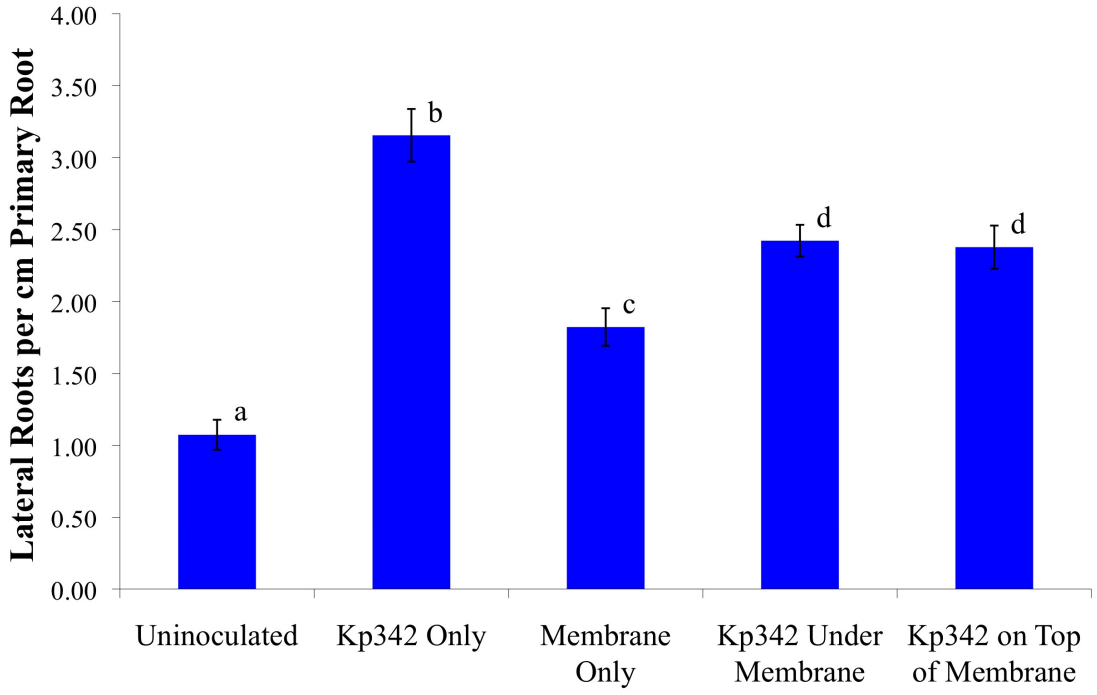


B

Figure 5-3. Bacterial strain specificity of lateral root promotion. A) Lateral root per cm primary root on Kp342 inoculated plants compared to plants inoculated with Kp13883 and *E. coli* K12. B) Lateral root per cm primary root on Kp342 inoculated plants compared to *E. cloacae* P101 inoculation. *Arabidopsis* plants were inoculated with 10 cells 9 days after germination and roots counted 9 days post inoculation. Statistical difference denoted by different letters.



A



B

Figure 5-4. Lateral root promotion due to a secreted product. A) Diagram of membrane experiment. B) Lateral roots per cm primary root on *Arabidopsis* with Kp342 inoculated directly on roots or with a 0.2 μm membrane between the roots and the bacteria. Controls include uninoculated plants with and without the membrane and Kp342 inoculated under the membrane. Plants were inoculated with 10 cells 9 days after germination and roots counted 9 days post inoculation. Statistical difference denoted by different letters.

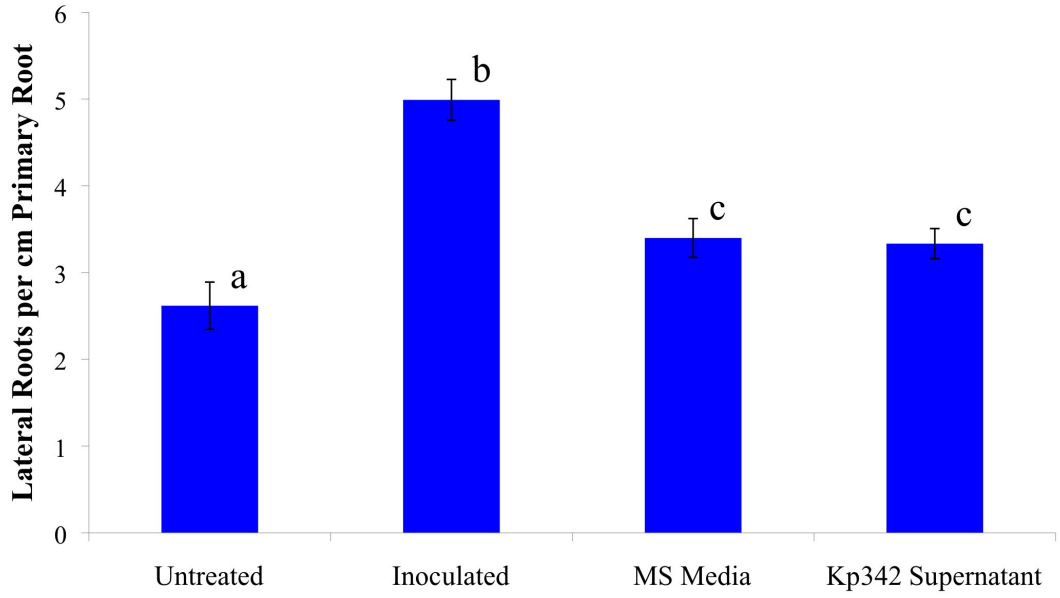


Figure 5-5. Lateral root promotion in response to culture supernatants. Lateral roots per cm primary root on *Arabidopsis* treated with concentrated Kp342 culture supernatant or uninoculated MS media. Kp342 inoculated and uninoculated plants used as positive and negative controls. Plants were treated 9 days after germination and roots counted 9 days post inoculation. Statistical difference denoted by different letters.

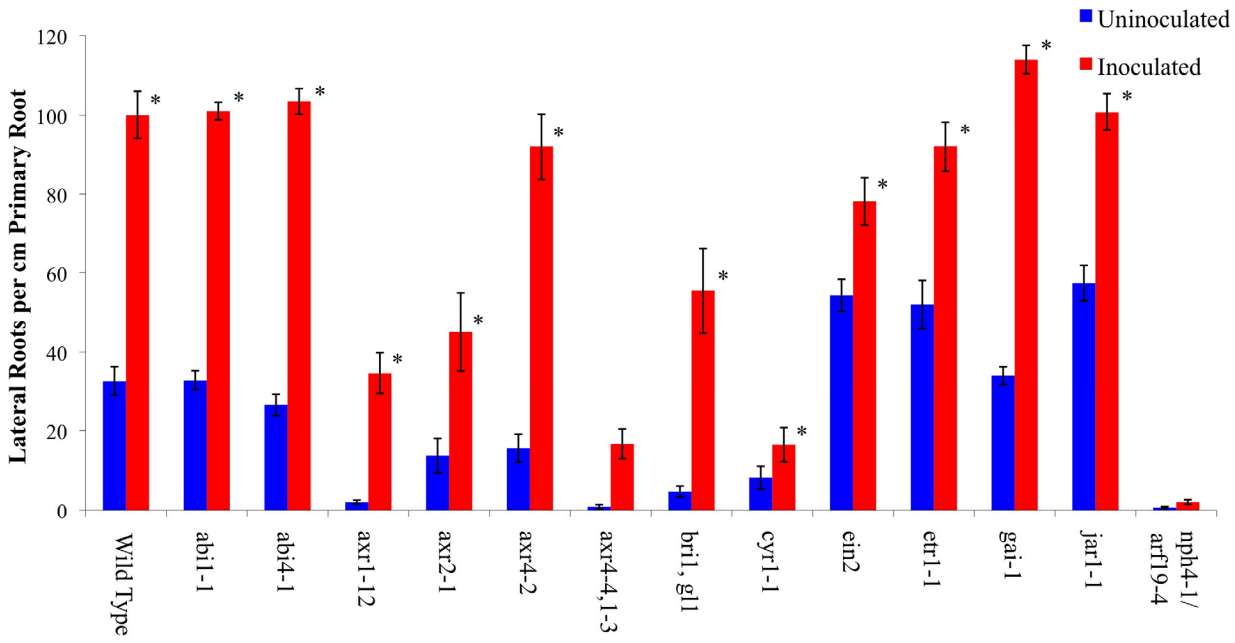
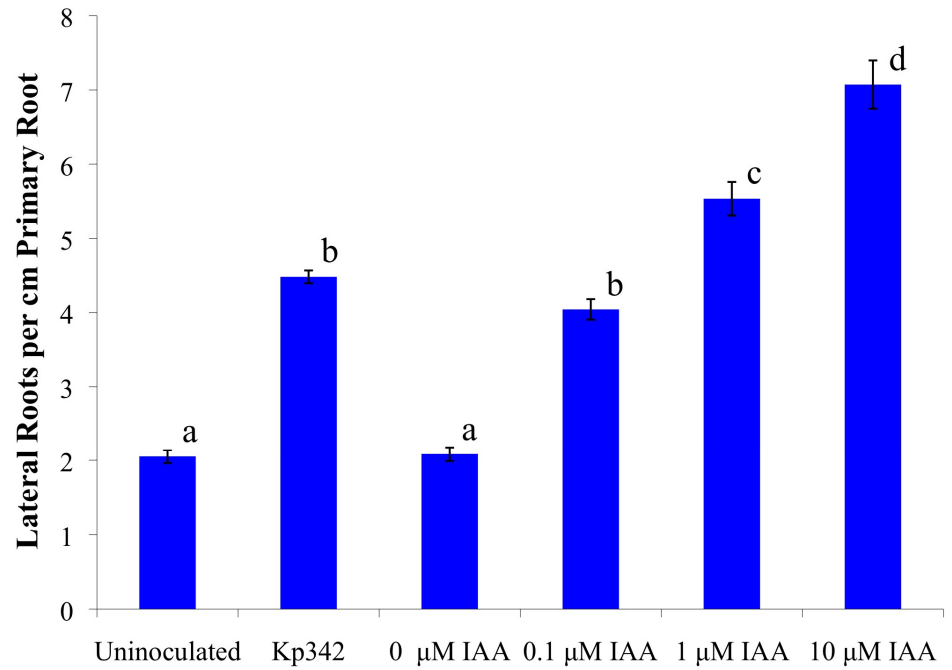
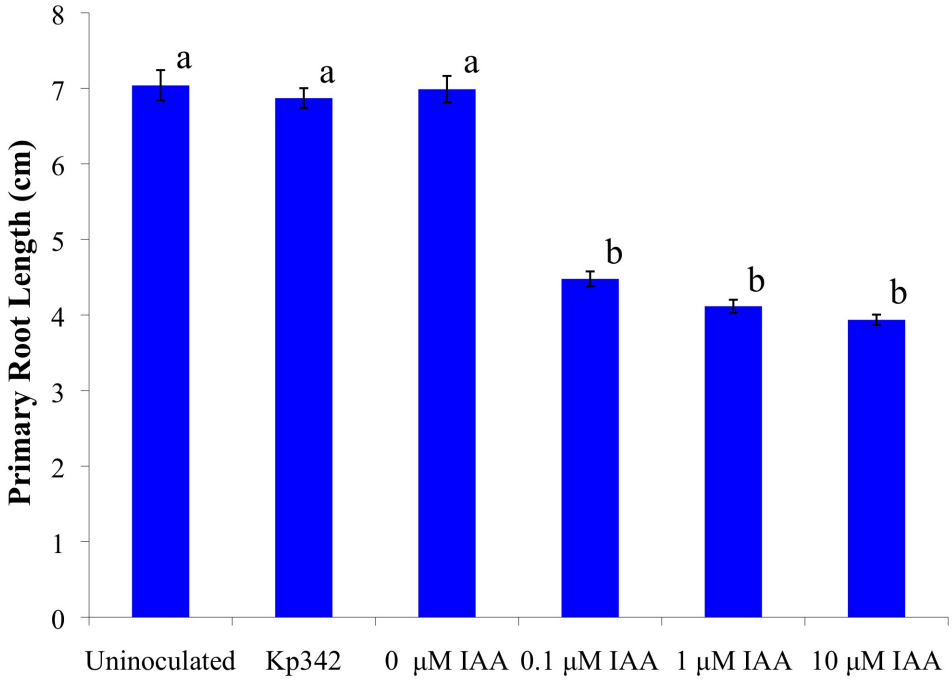


Figure 5-6. Response of hormone insensitive *Arabidopsis* mutants to Kp342. Lateral roots per cm primary root on uninoculated (blue) and Kp342 inoculated (red) plants. Plants were inoculated with 10 cells 9 days after germination and roots counted 9 days post inoculation. Asterisks indicate inoculated plants were statistically different from uninoculated controls.



A



B

Figure 5-7. Comparison of Kp342 inoculation to exogenous IAA treatment. A) Lateral roots per cm primary root and B) primary root length on Kp342 inoculated Col-0 *Arabidopsis* compared to plants grown on media supplemented with 0.1, 1, and 10 µM IAA. Plants were treated 9 days after germination and roots counted 9 days post inoculation. Statistical difference denoted by different letters.

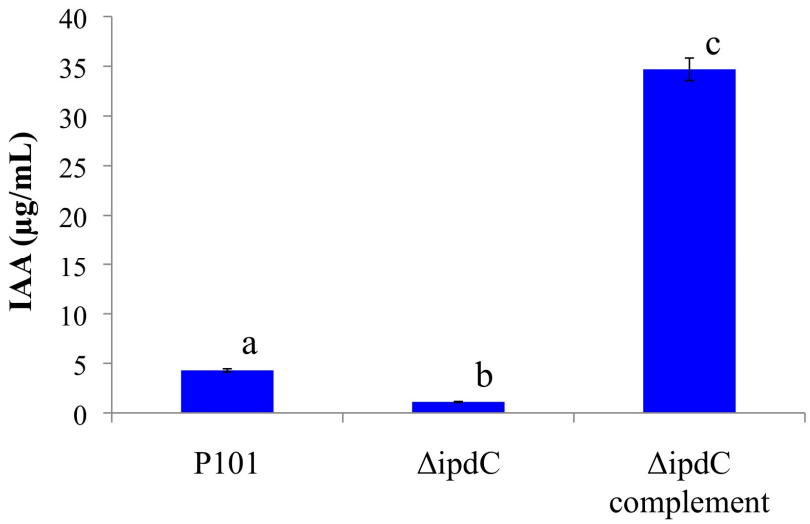


Figure 5-8. Auxin production in P101 strains. Concentration of indole-3-acetic acid in culture supernatants of P101, a Δ*ipdC* mutant, and the Δ*ipdC* mutant complemented with a copy of the *ipdC* gene carried on plasmid pHLT14. Bacteria were cultured in M9 media supplemented with 100 µg/mL tryptophan and supernatants were collected 46 hours after inoculation. IAA concentration adjusted for differences in culture OD_{620} in order to correct for differences in growth between cultures. Statistical difference denoted by different letters.

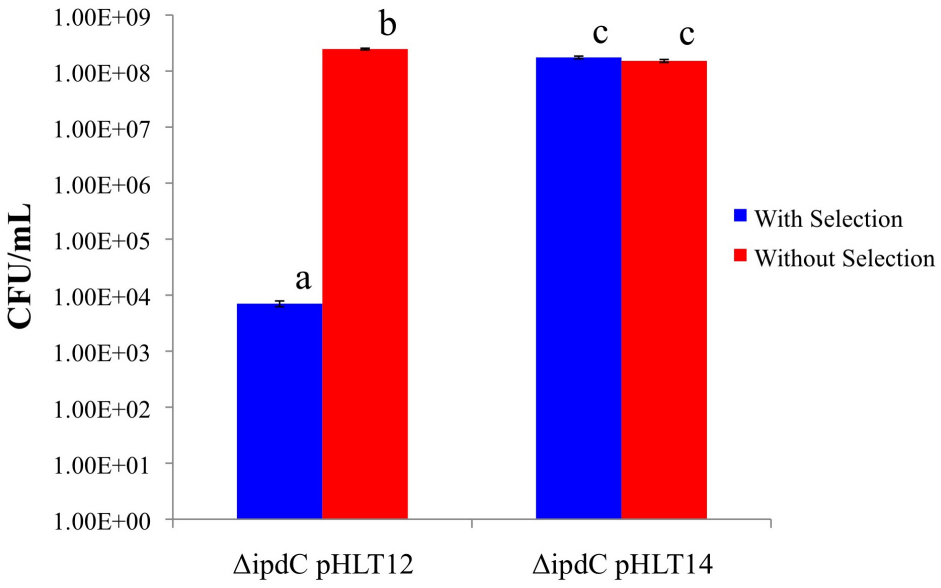


Figure 5-9. Plasmid stability in P101 Δ*ipdC*. Cultures were grown in LB without selection pressure. CFU/mL was counted after 2 subculturings by dilution spread-plating on media with and without antibiotics to determine how many cells carried the plasmid and were resistant. pHLT14 carries the *ipdC* complement with the *par* locus for plasmid stability. pHLT12 lacks the *par* locus and was used as a negative control.

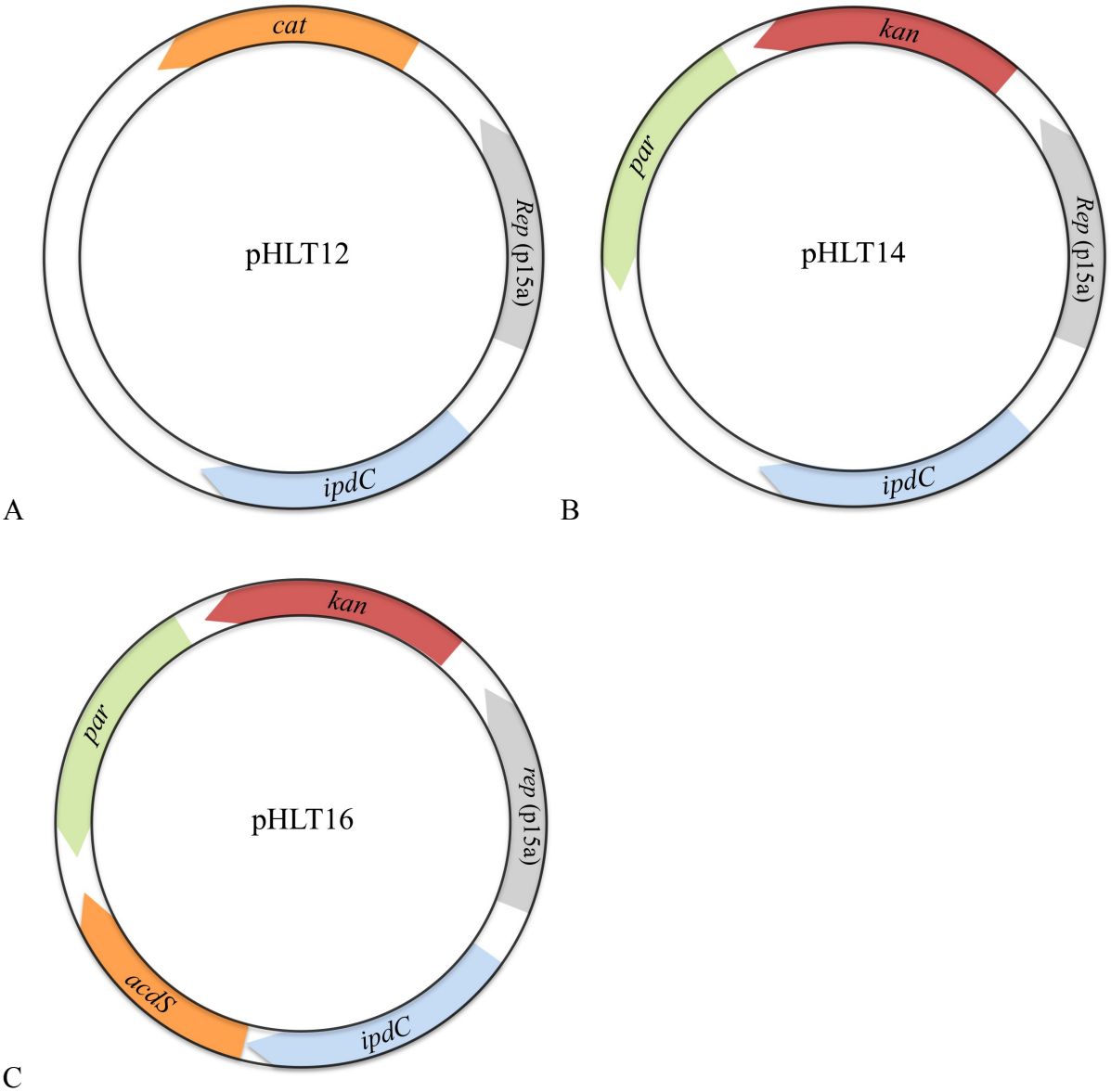


Figure 5-10. Diagrams of selected plasmid constructs. A) pHLT12 is a derivative of pACYC184 with the *ipdC* gene from *E. cloacae* P101 cloned into the BamHI resistriction site. B) pHLT14 is a derivative of pHLT12 with the *par* locus for plasmid stability from pTR102 and the kanamycin resistance gene from pCR4 cloned into the ScaI site. C) pHLT16 is a derivative of pHLT14 with the *acdS* gene for ACC deaminase from *P. putida* UW4 cloned immediately downstream of the *ipdC* gene.

Table 5-1. Examples of plant hormones secreted by bacteria

Bacteria	Hormone(s)	Reference(s)
Gluconacetobacter diazotrophicus	Auxin, Gibberellin	Bastian et al. 1998
Azospirillum sp.	Auxin, Cytokinin, Gibberellin	Crozier et al. 1988; Cacciari et al. 1989; Bottini et al. 1998
Azotobacter chroococcum	Cytokinin	Nieto and Frankenberger 1989
Bacillus sp.	Gibberellin	Gutierrez-Manero et al. 2001
Herbaspirillum seropedicae	Auxin, Gibberellin	Bastian et al. 1998
Klebsiella pneumoniae	Auxin	El-Khawas and Adachi 1999
Paenibacillus polymyxa	Auxin, Cytokinin	Lebuhn et al. 1997; Timmusk et al. 1999
Pseudomonas flourescens	Cytokinin	Garcia de Salamone et al. 2001

Table 5-2. Comparison of growth-promoting mechanisms in *K. pneumoniae* 342, *E. cloacae* P101, and *E. coli* K12

Gene product	*K. pneumoniae* 342	*E. cloacae* P101	*E. coli* K12
Indole-pyruvate decarboxylase (*ipdC*)	Present	Present	Absent
ACC deaminase	Absent	Absent	Putative
2,3-butanediol production (*budABC*)	Absent (only *budA*)	Present	Absent
Nitrogenase (*nif*)	Present (all 16 *nif* genes)	Absent (only *nifU*)	Absent (only *nifA* and *nifS*)
Phytase	Absent	Absent	Absent

Table 5-3. Hormone insensitive *Arabidopsis* mutants used to examine lateral root promotion

Mutant	Hormone(s)	Reference
axr1-12	Auxin	Lincoln et al. 1990
axr2-1	Auxin; Ethylene; Abscisic Acid	Wilson et al. 1990
axr4-2	Auxin	Hobbie and Estelle 1995
axr4-2,1-3	Auxin	Hobbie and Estelle 1995
bri1,gl	Brassinosteroid	Clouse et al. 1996
ein2	Ethylene	Alonso et al. 1999
etr1-1	Ethylene	Bleecker et al. 1988; Guzman and Ecker 1990
gai-1	Gibberellin	Koorneef et al. 1985
cyr1-1	Cytokinin	Deikman and Ulrich 1995
abi1-1	Abscisic Acid	Koornneef et al. 1984
abi4-1	Abscisic Acid	Finkelstein 1994
jar1-1	Jasmonic Acid	Staswick et al. 1992

Table 5-4. Gene deletion and plasmid construction primers

Primer	Sequence (5' – 3')
A_P101_ipdC	CGCGGATCCTATAACCCGCTTTGGTCGAG
D_P101_ipdC	CGCGGATCCGGATGATTGCGGTATGGTG
F2-JMP101ipdC	CATAACAGGTGTCCTTCTGAC
R2-JMP101ipdC	GCGCGTAATAACGCCTGATTA
par1sa	GGGGTCGACGACTACACCGAGGGGGAAAG
par2sa	CTGGTCGACGGGTAAGTACGCCATCAGGA
tpkanpar1	TCAGCTACTGGGCTATCTGGA
tpkanpar2	CTGGATCTCAACAGCGGTAAG
pACYCipdcRXhoI-1	CCTCGAGTCAGGCGTTATTACGCGCTT
pACYC184-FSbfI	ACTGGCCCTGCAGGTTTGAGAAGCACACGGTCAC
ACCd-FXhoI	GCTCGAGGATATCACAACAATCAAGGA
ACCd-RSbfI	CCTGTGCCTGCAGGACTGAATTGAACCCGAAACG

CHAPTER 6
CONCLUSION

Bacteria can interact with plants in a variety of different ways. Some act as pathogens while others live in association with plants without causing disease. Many plant-associated bacteria have beneficial effects on plants. Such bacteria can live as epiphytes on the surface of plant tissues, or they can live within plant tissue, as symbionts or endophytes. The common theme driving the research described in this work is the examination bacteria that inhabit the interior of plant tissues, whether they are a pathogen or growth promoter. This work also aimed to draw upon a wide range of research techniques and approaches, from utilizing bioinformatics with new high-throughput sequencing and optical mapping, to using classic molecular biology and plant culturing methods to answer questions about these plant-inhabiting bacteria.

The presence of bacteria in plant tissues is a common occurrence. In the case of the disease, Citrus Greening, also known as Huanglongbing (HLB), the nature of the microbial community within infected tissue is still open to debate. Since the bacterium believed to cause this disease, '*Candidatus* Liberibacter spp.', has yet to be isolated in pure culture, Koch's postulates have not been fulfilled and its role in disease development has not been confirmed. As a result, the presence and identity of any other bacteria within infected tissue have been investigated, since other pathogens may contribute to development of the disease. Different bacteria have been identified or cultured from HLB-infected plants in various citrus growing regions throughout the world. In China and Brazil, phytoplasmas have been identified in the midribs of leaves from infected citrus plants (Chen et al. 2009; Teixeira et al. 2008). In addition, Sagaram and associates (2009) reported a wide array of bacterial diversity in midribs from infected citrus plants in Florida, identifying 47 orders of bacteria in 15 different phyla using 16S ribosomal ribonucleic acid (rRNA) microarrays.

In the research described here, a metagenomic approach using high-throughput sequence data from HLB-infected tissue was employed to ascertain the microbial community within infected tissue by polymerase chain reaction (PCR) independent means, thus removing primer based biases from the analysis. In contrast to the other studies, '*Ca*. L. asiaticus' was the only bacterium identified in infected tissue. This difference is believed to be due to the nature of the tissues sampled in the different studies. The other studies examined leaf midribs whereas the research performed here looked solely at phloem tissue scraped from the inside of bark. Past studies on HLB-infected plants indicate the causative agent is limited to phloem sieve tubes, therefore, specifically focusing on phloem tissue was deemed the best choice for identifying bacteria associated with the disease. The midribs examined in the other studies would have contained leaf tissue, such as the apoplast, which is known to harbor bacterial endophytes. As a result, the presence of other bacteria in those tissues could vary based on other factors and may not play a role in the development of HLB. Given the large amount of sequence data obtained from the high-throughput methods used, extensive genome coverage of any bacterium present within infected phloem was expected, and '*Ca*. L. asiaticus' was the only bacterium with such coverage. As the genome sequence of '*Ca*. L. asiaticus' used in this study was isolated from the insect vector that transmits the disease, finding the same sequence in tissue from an infected citrus plant further substantiates the etiology of '*Ca* L. asiaticus' in HLB disease.

As previously mentioned, many bacteria that can live within plant tissue do not cause disease, and instead have beneficial effects on their plant host. Therefore, research on endophytic bacteria that live in plant tissue and increase plant growth was another main focus in this work. There are several bacterial endophytes that have been studied for their roles in plant growth promotion, including *Gluconacetobacter diazotrophicus* PAl 5, *Azoarcus* sp. BH72,

Herbaspirillum seropedicae Z67, and *Klebsiella pneumoniae* 342. The genomes of these strains have been or are currently being sequenced. In addition, genomes of close, non-endophytic relatives of these bacteria are also available. With the increasing availability of genomic information, comparative genomics can be used to identify genes unique to the endophytic and growth-promoting lifestyle.

In regards to the genome of *G. diazotrophicus* PAl 5, one of the first nitrogen fixing endophytic bacteria isolated, two different groups have sequenced this strain. Strangely, considerable differences were observed between the closed sequences reported by these two groups. Given this information, it appeared that one of these sequences might contain assembly or sequencing errors. Considering the genomic sequence of this strain will eventually be used for comparative genomics and the planning new experiments with the bacterium, it was important to determine which of these sequences is the best representation of the actual *G. diazotrophicus* PAl 5 genome. To that end, an optical map of the *G. diazotrophicus* PAl 5 chromosome was made in order to obtain a physical map to which the two sequences could be compared. From that comparison, one genomic sequence was found to have significant chromosomal rearrangements. In addition, examination of the annotation identified many differences in the predicted genes in the two sequences. Therefore, the underlying cause of the disparity between the two genomic sequences of the PAl 5 strain has resulted in significant differences at the gene level that could confound future studies. Possible explanations for the discordant PAl 5 sequence are that the bacterium sequenced has evolved in culture to the point that it is no longer the same as the original PAl 5 strain or that errors occurred during the sequencing and assembly. This situation surrounding the two PAl 5 genome sequences demonstrates the importance of using optical maps when assembling genomes and providing raw

sequencing reads with quality scores when submitting genomes to public databases. With this information, independent verification can be made when contradictory sequences arise and it can be determined if the differences are due to strain evolution or sequencing errors. In either case, the analysis performed in this work demonstrates which of the two genomic sequences should be used to guide future research on *G. diazotrophicus* PAl 5.

New high-throughput techniques allowing examination of plant-associated bacteria at the whole genome level has opened up avenues for comparative genomics, but the ultimate goal of these techniques is still to identify gene targets in the bacteria to examine how they colonize plants and increase their growth. There is only so much information that sequence analysis can reveal about an organism without performing laboratory experiments to confirm the roles of these genes in plant growth and colonization. For this reason, another focus of the research described here involved working directly with endophytes, looking at their effect on root growth in *Arabidopsis* as an assay for studying their plant growth-promoting effects. A variety of experiments were performed on the root growth phenotype. Because genomes of the two endophytic bacteria examined were found to possess a gene for auxin synthesis, this gene was selected as a target for studying growth promotion by these strains. While unexpected difficulties arose in the reproducibility of the lateral root promoting phenotype, a knockout and complementing plasmids of the auxin synthesis gene were constructed in the endophyte, *E. cloacae* P101. Therefore, genetic techniques and vectors for studying the roles of other genes of interest in P101 are available for future studies.

The other endophytic bacterium studied in this research, *K. pneumoniae* 342 (Kp342), was not amenable to genetic manipulation. In addition, examination of its genome revealed the presence of numerous pathogenicity and antibiotic resistance genes (Fouts et al. 2008).

Subsequent benchwork on this bacterium confirmed the antibiotic resistance and pathogenicity of this strain. Therefore, Kp342 is not a suitable candidate for agricultural application, though it may be useful as an experimental endophytic model should a convenient system for making in-frame deletions ever be developed to work reliably in this strain. In the current study, given its pathogenic nature, it was deemed inefficient to expend more time and effort on developing such a system for Kp342. As a result, subsequent studies focused on another, more genetically tractable endophyte.

In closing, the research described here encompasses work done on a variety of different plant-associated bacteria. In addition, different techniques were employed to examine plant-microbe interactions over a range of depths. The ultimate goal of this research was to utilize different techniques and technologies for an integrated approach to examining plant-microbe interactions, from genomics and metagenomics to the examination of specific genes identified in plant-associated bacteria.

APPENDIX A
GENOME OF *KLEBSIELLA PNEUMONIAE* 342

In order to direct and facilitate future studies on *Klebsiella pneumoniae* 342 (Kp342), the genome of this bacterium was sequenced. With the genome sequence in hand, Kp342 was compared to other plant-associated bacteria in order to identify potential plant colonization and growth promotion genes. Given its plant origins and its relation to other human pathogens, the genome was also examined to determine its pathogenic potential. The Kp342 genome was examined for the presence of antibiotic resistance genes and pathogenicity genes in other *K. pneumoniae* strains. The ultimate goal of the genome sequencing project was to identify targets for gene knockout once a genetic system for K342 is developed. The following work includes text excerpts, data, and analyses that were contributed to the Kp342 genome paper (Fouts et al. 2008). Bioinformatic analyses were performed by Derric Fouts. Pathogenicity assays were performed by Carston Struve. Heather Tyler wrote about the biological significance of plant-induced genes identified, performed antibiotic resistance assays on Kp342, and performed statistical analysis on the pathogenicity data.

Plant Induced Genes Found in Kp342

Bioinformatic analysis of the Kp342 genome performed by Derrick Fouts demonstrated that this endophyte possesses many genes that share homology with known plant-induced genes. Several of these genes were selected as targets for future mutagenesis to confirm their role in endophytic colonization by Kp342. Of these genes, many appear to be involved in the adaptation of bacteria to conditions within plant tissue, such as limited amino acid concentration and carbon sources. Several amino acid and nucleotide biosynthesis genes present in Kp342 were found to be induced in *Ralstonia solanacearum* and *Pseudomonas syringae* pv. tomato upon plant colonization. These genes include CTP synthase (*pyrG*), acetyl-CoA

acetyltransferase, amidophosphoribosyltransferase (*purF*), argininosuccinate synthase (*argG*), a diaminopimelate decarboxylase (*lysA*), and acetolactate synthase large subunit (*ivlI*) (Boch et al. 2002; Brown and Allen 2004). The importance of amino acid biosynthesis in plant-microbe interactions is supported by the observation that *P. syringae* mutants impaired in the biosynthesis of some amino acids were unable to cause disease symptoms on tomato (Cuppels 1986). Further gene deletion studies in Kp342 would be able to confirm if these genes also play a role in endophytic adaptation to plant tissue in addition to their role in plant pathogens.

Other plant-inducible genes that Kp342 shares homology with are thought to be involved in adapting to plant defenses. Putative stress response genes turned on in *R. solanacearum* upon plant colonization also found in Kp342 include an Ada regulatory protein, excinuclease (*uvrA*), deoxyribonucleic acid (DNA)-damage-inducible protein F (*dinF*), fumerate hydratase (*fumC*), and an acriflavin resistance protein A (*acrA*) (Brown and Allen 2004). Several of these genes have been implicated in bacterial response to oxidative stress and DNA damage due to plant defense responses, some of which are involved in DNA repair. For example, the Ada protein is required to activate the transcription of genes involved in adaptive response to DNA methylation damage caused by alkylating agents, and has also been shown to be activated by nitric oxide (Landini and Volkert 1995; Nakabeppu and Sekiguchi 1986; Vasilieva and Maschkovskaya 2005). In addition, exinuclease (*uvrA*) functions in ultraviolet (UV) induced DNA repair, but has also been shown to participate in hydrogen peroxide and toxic chemical induced DNA damage repair, indicating that this gene may act to protect the bacteria against DNA-damaging compounds produced by plants (Asad et al. 1994; Mikulasova et al. 2005; Rupp et al. 1982). These oxidative response genes are not limited to DNA repair pathways. Fumarate hydratase (*fumC*), which is synthesized highest under conditions when superoxide radicals accumulate, is

part of the tricarboxylic acid cycle (Park and Gunsalus 1995). This gene appears to be a back up for the more abundant FumA, which is inactivated under oxidative conditions (Park and Gunsalus 1995; Ueda et al. 1991). Since an early plant defense response involves the increase of reactive oxygen species, induction of oxidative stress related genes indicate the bacteria are actively evading this defense mechanism while colonizing plants. Acriflavine resistance protein A (*acrA*) is another stress response gene induced upon plant colonization, but does not appear to be triggered by oxidative stress. The product of this gene encodes a component of the AcrAB-TolC efflux pump that is important in toxic waste removal in bacteria and shows increased expression under stress conditions (Helling et al. 2002; Ma et al. 1995). In the plant pathogen, *Erwinia amylovora*, AcrAB is induced by phytoalexin plant defense compounds from apple and is required for resistance to these toxins (Burse et al. 2004). These genes were identified as being induced in plant pathogens, so their role in an endophytic bacterium's response to plant defense should be confirmed.

A gene believed to be involved in plant attachment has also been identified for further study. This plant-inducible haemagglutinin gene in *R. solacacearum* is homologous to a Kp342 filamentous haemagglutinin (HecA) protein (Brown and Allen 2004). The *hecA* gene is part of a HecA/B hemolysin/hemagglutinin secretion operon. The HecA/B proteins make up a two-partner secretion (TPS) system in which a TpsA family exoprotein with specific conserved secretion signals is transported across the membrane by a TpsB family channel-forming transporter that recognizes the secretion signal (Jacob-Dubuisson et al. 2001). In *Erwinia chrysanthemi*, a mutant in the *hecA* gene that encodes an adhesin had reduced attachment, cell aggregate formation, and virulence on *Nicotinia clevelandii* (Rojas et al. 2002). Homologs of this gene appear in both plant and animal pathogens (Rojas et al. 2002). A TPS operon (*hlpAB*)

has also been identified in the nonpathogenic plant colonizing bacteria, *Pseudomonas putida* KT2440, and is necessary for competitive root colonization (Molina et al. 2006). The presence of another TPS operon important to colonization in a nonpathogenic plant-associated bacteria gives support the likelihood that the HecA/B homologue in Kp342 is a promising candidate for future study.

Antibiotic Resistance in Kp342

Annotation of the Kp342 genome also revealed the presence of several antibiotic resistance genes. The majority of these genes are efflux pumps and beta-lactamase genes that confer resistance to a wide array of antibiotics. Resistance of Kp342 to antibiotics conferred by these genes was further confirmed in the lab by the Kirby-Baur disc method, testing members of all major antibiotic families (Table A-1). Since Kp342 is a plant isolate as opposed to a clinical isolate, the presence of so many broad-spectrum antibiotic resistances in this strain was intriguing. One possible explanation for the maintenance of such resistance genes in Kp342 is for the removal of toxic metabolites produced by plants. Another possible explanation for the antibiotic resistances in Kp342 is exposure to the plant signaling molecule, salicylic acid. Salicylate is known to enhance antibiotic resistance when co-applied to bacterial cultures of *Serratia marcescens* and *Escherichia coli* (Berlanga and Viñas 2000; Cohen et al. 1993). It has also been found to increase many antibiotic resistances in clinical *K. pneumoniae* strains, including resistance to cefazolin, cefoperazone, norfloxacin, doxycycline, mezlocillin and trimethoprim-sulphamethoxasole (Domenico et al. 1990). Because salicylic acid is pivotal in signaling many plant processes, such as responses to abiotic stress and defense responses to pathogens (Raskin 1992), Kp342 would have been exposed to this hormone during its endophytic lifestyle inside plant tissue. As a result, long-term exposure to this hormone within

the plant could have contributed to the maintenance of so many antibiotic resistance genes in the genome.

Kp342 Pathogenicity

Given the number of antibiotic resistance genes found in Kp342 and the fact that other strains of *K. pneumoniae* are known to be pathogenic to humans, the virulence and pathogenicity of Kp342 needed to be investigated before considering its use in an agricultural setting. Several virulence factors in pathogenic *K. pneumoniae* strains have been identified by attenuation of disease in signature-tagged mutagenesis studies (Lawlor et al. 2005; Struve et al. 2003; Lai et al. 2001). Therefore, these genes were searched for in the Kp342 genome. Several of these pathogenicity genes were identified in the Kp342 genome, and with the identification of known virulence genes in Kp342, it was deemed prudent to verify the level of pathogenicity of this endophyte in an animal model.

In order to determine pathogenicity in animals, Kp342 was assayed in urinary tract and lung infection models in mice by a collaborator, Carsten Struve, at the Statens Serum Institute in Copenhagen, Denmark. For comparison, another clinical isolate of *K. pneumoniae*, C3091, was included in the study. In the urinary tract model, five out of six mice had infected bladders three days after inoculation with Kp342 in numbers similar to those seen when inoculated by the clinical isolate (Table A-2). Kp342 was also able to ascend from the bladder to the kidneys, but at a level 28 times lower than the clinical strain. In the lung infection model, all Kp342 inoculated mice were infected two days after inhalation, but infection was 49 times lower than in mice inoculated with C3091. When looking at systemic spreading of Kp342 from the lungs, only one in five mice had infections that spread to the liver and none of the mice had infected spleens (Table A-2). In contrast, three and two of the five C3091 inoculated mice had infected

livers and spleens, respectively (Table A-2). Therefore, while Kp342 has the potential to be pathogenic and is capable of causing infection in an animal host, its effects are attenuated and less virulent compared to clinical isolates.

Table A-1. Kp342 Antibiotic Resistance Profile

Drug (μg)	Phenotype	Observed zone of inhibition diameter (mm)[a]	Interpretive standards (mm)[a]
Novobiocin (30)	Resistant	0	≤17 resistant
Gentamicin (10)	Intermediate	14	13-14 intermediate
Kanamycin (30)	Sensitive	30	≥18 sensitive
Neomycin (30)	Resistant	9	≤12 resistant
Cefotaxime (30)	Sensitive	25	≥23 sensitive
Cefoperazone (75)	Sensitive	24	≥21 sensitive
Cefazolin (30)	Sensitive	20	≥18 sensitive
Ceftriaxone (30)	Intermediate	15	14-20 intermediate
Cefuroxime (30)	Intermediate	15	15-17 intermediate
Cephalothin (30)	Resistant	14	≤14 resistant
Moxalactam (30)	Intermediate	22	15-22 intermediate
Ampicillin (10)	Resistant	0	≤15 resistant
Mezlocillin (75)	Intermediate	18	18-20 intermediate
Penicillin (10)[b]	Resistant	0	
Piperacillin (100)	Intermediate	19	18-20 intermediate
Ticarcillin (75)	Resistant	7	≤14 resistant
Azithromycin (15)[b]	Undetermined	10	
Erythromycin (15)[b]	Resistant	0	
Ciproflaxacin (5)	Intermediate	18	16-20 intermediate
Nalidixic acid (30)	Resistant	8	≤13 resistant
Norfloxacin (10)	Resistant	0	≤12 resistant
Oxolinic acid (2)	Resistant	7	≤10 resistant
Sulfisoxazole (0.25)	Resistant	8	≤12 resistant
Trimethoprim (5)	Resistant	0	≤10 resistant
Minocycline (30)	Resistant	8	≤14 resistant
Oxytetracycline (30)[b]	Resistant	0	
Tetracycline (30)	Resistant	10	≤14 resistant
Rifampin (5)[b]	Resistant	0	
Spectinomycin (100)[b]	Resistant	0	

[a]Observed zones of inhibition were compared to interpretive standards reported for members of the Enterobacteriaceae. [b]No interpretive standards for Enterobacteriaceae were reported.

Table A-2. Infection of *K. pneumoniae* 342 and C3091 in a Mouse Infection Model

Model	Tissue	Log CFU Kp342[a]	Log CFU C3091[a]
Urinary track infection	Bladder	3.40 +/- 0.72	3.94 +/- 0.40
Urinary track infection	Kidney[b]	2.43 +/- 0.40	3.87 +/- 0.45
Lung infection	Liver	0.44 +/- 0.44	1.99 +/- 1.02
Lung infection	Lung[b]	4.63 +/- 0.41	6.32 +/- 0.38
Lung infection	Spleen	0	0.54 +/- 0.35

[a]Mean of the log of colony forming units (CFU) recovered per organ plus or minus the standard error. [b]Tissues had statistically significant difference between Kp342 and C3091 infection at the 5% level as determined by Fisher's Least Significant Difference Test.

LITERATURE CITED

Abeles, F. B., Morgan, P. W., and Salveit, M. E. Jr. 1992. The biosynthesis of ethylene. Pages 26-55 in: Ethylene in Plant Biology, 2nd ed. Academic Press, San Diego, CA.

Alonso, J. M., Hirayama, T., Roman, G., Nourizadeh, S., and Ecker, J. R. 1999. EIN2, a bifunctional transducer of ethylene and stress responses in *Arabidopsis*. Science 284:2148-2152.

Anderson, D. M., and Mills, D. 1985. The use of transposon mutagenesis in the isolation of nutritional and virulence mutants in 2 pathovars of *Pseudomonas syringae*. Phytopathology 75:104-108.

Armon, R., Dosoretz, C. G., Azov, Y., and Shelef, G. 1994. Residual contamination of crops irrigated with effluent of different qualities-a field-study. Water Sci. Technol. 30:239-248.

Asad, L. M. B. O., Dealmeida, C. E. B., Dasilva, A. B., Asad, N. R., and Leitao, A. C. 1994. Hydrogen-peroxide induces the repair of UV-damages DNA in *Escherichia coli* – A LexA-independent but UvrA-dependent and RecA-dependent mechanism. Curr. Microbiol. 29:291-294.

Aston, C., Mishra, B., and Schwartz, D. C. 1999. Optical mapping and its potential for large-scale sequencing projects. Trends Biotechnol. 17:297-302.

Auty, M., Duffy, G., O'Beirne, D., McGovern, A., Gleeson, E., and Jordan, K. 2005. In situ localization of *Escherichia coli* O157: H7 in food by confocal scanning laser microscopy. J. Food Prot. 68:482-486.

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., and Zagnitko, O. 2008. The RAST Server: Rapid annotations using subsystems technology. BMC Genomics 9:75.

Babic, I., Roy, S., Watada, A. E., and Wergin, W. P. 1996. Changes in microbial populations on fresh cut spinach. Int. J. Food Microbiol. 31:107-119.

Barak, J. D., Gorski, L., Naraghi-Arani, P., and Charkowski, A. O. 2005. *Salmonella enterica* virulence genes are required for bacterial attachment to plant tissue. Appl. Environ. Microbiol. 71:5685-5691.

Barak, J. D., Whitehand, L. C., and Charkowski, A. O. 2002. Differences in attachment of *Salmonella enterica* serovars and *Escherichia coli* O157: H7 to alfalfa sprouts. Appl. Environ. Microbiol. 68:4758-4763.

Barbieri, P., and Galli, E. 1993. Effect on wheat root development of inoculation with an *Azospirillum brasilense* mutant with altered indole-3-acetic acid production. Res. Microbiol. 144:69-75.

Bardwell, J. C. A., McGovern, K., and Beckwith, J. 1991. Identification of a protein required for disulfide bond formation in vivo. Cell 67:581-589.

Bartel, B. 1997. Auxin Biosynthesis. Annu. Rev. Plant Physiol. Plant Mol. Biol. 48:51-66.

Bastian, F., Cohen, A., Piccoli, P., Luna, V., Baraldi, R., and Bottini, R. 1998. Production of indole-3-acetic acid and gibberellins A_1 and A_3 by *Acetobacter diazotrophicus* and *Herbaspirillum seropedicae* in chemically-defined culture media. Plant Growth Regul. 24:7-11.

Berlanga, M., and Viñas, M. 2000. Salicylate induction of phenotypic resistance to quinolones in *Serratia marcescens*. J. Antimicrob. Chemother. 46:279-282.

Bernier, S. P., Silo-Suh, L., Woods, D. E., Ohman, D. E., and Sokol, P. A. 2003. Comparative analysis of plant and animal models for characterization of *Burkholderia cepacia* virulence. Infect. Immun. 71:5306-5313.

Bernstein, N., Sela, S., Pinto, R., and Ioffe, M. 2007. Evidence for internalization of *Escherichia coli* into the aerial parts of maize via the root system. J. Food Prot. 70:471-475.

Besser, T. E., LeJeune, J. T., Rice, D. H., Berg, J., Stilborn, R. P., Kaya, K., Bae, W., and Hancock, D. D. 2005. Increasing prevalence of *Campylobacter jejuni* in feedlot cattle through the feeding period. Appl. Environ. Microbiol. 71:5752-5758.

Bleecker, A. B., Estelle, M. A., Somerville, C., and Kende, H. 1988. Insensitivity to ethylene conferred by a dominant mutation in *Arabidopsis thaliana*. Science 241:1086-1089.

Boch, J., Joardar, V., Gao, L., Robertson, T. L., Lim, M., and Kunkel, B. N. 2002. Identification of *Pseudomonas syringae* pv. tomato genes induced during infection of *Arabidopsis thaliana*. Mol. Microbiol. 44:73-88.

Boddey, R. M. 1995. Biological nitrogen-fixation in sugar-cane - a key to energetically viable biofuel production. Crit. Rev. Plant Sci. 14:263-279.

Boddey, R. M., de Oliveira, O. C., Urquiaga, S., Reis, V. M., de Olivares, F. L., Baldani, V. L. D., and Döbereiner, J. 1995. Biological nitrogen-fixation associated with sugar cane and rice -contributions and prospects for improvement. Plant Soil 174:195-209.

Bohm, M., Hurek, T., and Reinhold-Hurek, B. 2007. Twitching motility is essential for endophytic rice colonization by the N_2-fixing endophyte *Azoarcus* sp. strain BH72. Mol. Plant-Microbe Interact. 20:526-533.

Bosshard, P. P., Stettler, R., and Bachofen, R. 2000. Seasonal and spatial community dynamics in the meromictic Lake Cadagno. Arch. Microbiol. 174:168-174.

Bottini, R., Fulchieri, M., Pearce, D., and Pharis, R. P. 1989. Identification of gibberellin-A_1, gibberellin-A_3, and iso-A_3 in cultures of *Azospirillum lipoferum*. Plant Physiol. 90:45-47.

Bové, J. M. 2006. Huanglongbing: A destructive, newly-emerging, century-old disease of citrus. J. Plant. Pathol. 88:7-37.

Brandl, M. T. 2006. Fitness of human enteric pathogens on plants and implications for food safety. Annu. Rev. Phytopathol. 44:367-392.

Brandl, M.T., and Mandrell, R. E. 2002. Fitness of *Salmonella enterica* serovar Thompson in the cilantro phyllosphere. Appl. Environ. Microbiol. 68:3614-3621.

Brown, D. G. and Allen, C. 2004. *Ralstonia solanacearum* genes induced during growth in tomato: an inside view of bacterial wilt. Mol. Microbiol. 53:1641-1660.

Brüggemann, H., Henne, A., Hoster, F., Liesegang, H., Wiezer, A., Strittmatter, A., Hujer, S., Dürre, P., and Gottschalk, G. 2004. The complete genome sequence of *Propionibacterium acnes*, a commensal of human skin. Science 305:671-673.

Burnett, S. L., Chen, J. R., and Beuchat, L. R. 2000. Attachment of *Escherichia coli* O157: H7 to the surfaces and internal structures of apples as detected by confocal scanning laser microscopy. Appl. Environ. Microbiol. 66:4679-4687.

Burse, A., Weingart, H., and Ullrish, M. S. 2004. The phytoalexin-inducible multidrug efflux pump AcrAB contributes to virulence in the fire blight pathogen, *Erwinia amylovora*. Mol. Plant-Microbe Interact. 17:43-54.

Cacciari, I., Lippi, D., Pietrosanti, T., and Pietrosanti, W. 1989. Phytohormone-like substances produced by single and mixed diazotrophic cultures of *Azospirillum* and *Arthrobacter*. Plant Soil 115:151-153.

Calvin, L. 2007. Outbreak linked to spinach forces reassessment of food safety practices. Amber Waves 5:24-31.

Capoor, S. P., Rao, D. G., and Viswanat, S. M. 1967. *Diaphorina citri* Kuway., a vector of greening disease of citrus in India. Indian J. Agr. Sci. 37:572-576.

Carreno-Lopez, R., Campos-Reales, N., Elmerich, C., and Baca, B. E. 2000. Physiological evidence for differently regulated tryptophan-dependent pathways for indole-3-acetic acid synthesis in *Azospirillum brasilense*. Mol. Gen. Genet. 264:521-530.

Casimiro, I., Beeckman, T., Graham, N., Bhalerao, R., Zhang, H., Casero, P., Sandberg, G., and Bennett, M. J. 2003. Dissecting *Arabidopsis* lateral root development. Trends Plant Sci. 8:165-171.

Chang, A. C. Y., and Cohen, S. N. 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. 134:1141-1156.

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

Chelius, M. K., and Triplett, E. W. 2000. Immunolocalization of dinitrogenase reductase produced by *Klebsiella pneumoniae* in association with *Zea mays* L. Appl. Environ. Microbiol. 66:783-787.

Chelius, M. K., and Triplett, E. W. 2001. The diversity of archaea and bacteria in association with the roots of *Zea mays* L. Microb. Ecol. 41:252-263.

Chen, J., Pu, X., Deng, X., Liu, S., Li, H., and Civerolo, E. 2009. A phytoplasma related to '*Candidatus* Phytoplasma asteri' detected in citrus showing Huanglongbing (Yellow Shoot Disease) symptoms in Guangdong, P. R. China. Phytopathology 99:236-242.

Chen, Q., Savarino, S. J., and Venkatesan, M. M. 2006. Subtractive hybridization and optical mapping of the enterotoxigenic *Escherichia coli* H10407 chromosome: isolation of unique sequences and demonstration of significant similarity to the chromosome of *E. coli* K-12. Microbiology 152:1041-1054.

Chevalier, F., Pata, M., Nacry, P., Doumas, P., and Rossignol, M. 2003. Effects of phosphate availability on root system architecture: large-scale analysis of the natural variation between *Arabidopsis* accessions. Plant Cell Environ. 26:1839-1850.

Clouse, S. D., Langford, M., and McMorris, T. C. 1996. A brassinosteroid-insensitive mutant in *Arabidopsis thaliana* exhibits multiple defects in growth and development. Plant Physiol. 11:671-678.

Coenye, T., Vandamme, P., Govan, J. R. W., and Lipuma, J. J. 2001. Taxonomy and identification of the *Burkholderia cepacia* complex. J. Clin. Microbiol. 39:3427-3436.

Cohen, S. P., Levy, S. B., Foulds, J., and Rosner, J. L. 1993. Salicylate induction of antibiotic resistance in *Escherichia coli*: activation of the *mar* operon and a *mar*-independent pathway. J. Bacteriol. 175:7856-7862.

Cooley, M. B , Chao, D., and Mandrell, R. E. 2006. *Escherichia coli* O157: H7 survival and growth on lettuce is altered by the presence of epiphytic bacteria. J. Food Prot. 69:2329-2335.

Cooley, M. B., Miller, W. G., and Mandrell, R. E. 2003. Colonization of *Arabidopsis thaliana* with *Salmonella enterica* and enterohemorrhagic *Escherichia coli* O157: H7 and competition by *Enterobacter asburiae*. Appl. Environ. Microbiol. 69:4915-4926.

Cox-Foster, D. L., Conlan, S., Holmes, E. C., Palacios, G., Evans, J. D., Moran, N. A., Quan, P. L., Briese, T., Hornig, M., Geiser, D. M., Martinson, V., vanEngelsdorp, D., Kalkstein, A. L., Drysdale, A., Hui, J., Zhai, J. H., Cui, L. W., Hutchison, S. K., Simons, J. F., Egholm, M., Pettis, J. S., and Lipkin, W. I. 2007. A metagenomic survey of microbes in honey bee colony collapse disorder. Science 318:283-287.

Crozier, A., Arruda, P., Jasmim, J. M., Monteiro A. M., and Sandberg, G. 1988. Analysis of indole-3-acetic acid and related indoles in culture medium from *Azospirillum lipoferum* and *Azospirillum brasilense*. Appl. Environ. Microbiol. 54:2833-2837.

Cuppels, D. A. 1986. Generation and characterization of Tn5 insertion mutations in *Pseudomonas syringae* pv. *tomato*. Appl. Environ. Microbiol. 51:323-327.

Davis, M. J., Mondal, S. N., Chen, H. Q., Rogers, M. E., and Brlansky, R. H. 2008. Co-cultivation of '*Candidatus* Liberibacter asiaticus' with Actinobacteria from Citrus with Huanglongbing. Plant Dis. 92:1547-1550.

De Leij, F. A. A. M., Dixon-Hardy, J. E., and Lynch, J. M. 2002. Effect of 2,4-diacetylphloroglucinol-producing and non-producing strains of *Pseudomonas fluorescens* on root development of pea seedlings in three different soil types and its effect on nodulation by *Rhizobium*. Biol. Fertil. Soils. 35:114-121.

Debi, B. R., Taketa, S., and Ichii, M. 2005. Cytokinin inhibits lateral root initiation but stimulates lateral root elongation in rice (*Oryza sativa*). J. Plant Physiol. 162:507-515.

Deikman, J., and Ulrich, M. 1995. A novel cytokinin-resistant mutant of *Arabidopsis* with abbreviated shoot development. Planta 195:440-449.

Dobbelaere, S., Croonenborghs, A., Thys, A., Vande Broek, A., and Vanderleyden, J. 1999. Phytostimulatory effect of *Azospirillum brasilense* wild type and mutant strains altered in IAA production on wheat. Plant Soil 212:155-164.

Domenico, P., Hopkins, T., and Cunha, B. A. 1990. The effect of sodium salicylate on antibiotic susceptibility and synergy in *Klebsiella pneumoniae*. J. Antimicrob. Chemother. 26:343-351.

Dong, Y. M., Iniguez, A. L., Ahmer, B. M. M., and Triplett, E. W. 2003a. Kinetics and strain specificity of rhizosphere and endophytic colonization by enteric bacteria on seedlings of *Medicago sativa* and *Medicago truncatula*. Appl. Environ. Microbiol. 69:1783-1790.

Dong, Y. M., Iniguez, A. L., and Triplett, E. W. 2003b. Quantitative assessments of the host range and strain specificity of endophytic colonization by *Klebsiella pneumoniae* 342. Plant Soil 257:49-59.

Doddapaneni, H., Liao, H., Lin, H., Bai, X., Zhao, X., Civerolo, E. L., Irey, M., Coletta-Filho, H., and Pietersen, G. 2008. Comparative phylogenomics and multi-gene cluster analyses of the Citrus Huanglongbing (HLB)-associated bacterium *Candidatus* Liberibacter. BMC Res. Notes 1:72.

Dorr, J., Hurek, T., and Reinhold-Hurek, B. 1998. Type IV pili are involved in plant-microbe and fungus-microbe interactions. Mol. Microbiol. 30:7-17.

Duan, Y., Zhou, L., Hall, D. G., Li, W., Doddapaneni, H., Lin, H., Liu, L., Vahling, M., Gabriel, D. W., Williams, K. P., Dickerman, A., Sun, Y., and Gottwald, T. 2009. Complete genome sequence of citrus Huanglongbing bacterium, '*Candidatus* Liberibacter asiaticus' obtained through metagenomics. Mol. Plant Microbe Interact. 22(8):1011-1020.

Duarte, X., Anderson, C. T., Grimson, M., Barabote, R. D., Strauss, R. E., Gollahon, L. S. and San Francisco, M. J. D. 2000. *Erwinia chrysanthemi* strains cause death of human gastrointestinal cells in culture and express an intimin-like protein. FEMS Microbiol. Lett. 190:81-86.

Dubrovsky, J. G., Doerner, P. W, Colon-Carmona, A., and Rost, T. L. 2000. Pericycle cell proliferation and lateral root initiation in Arabidopsis. Plant Physiol. 124:1648-1657.

Eblen, B. S., Walderhaug, M. O., Edelson-Mammel, S., Chirtel, S. J., De Jesus, A., Merker, R. I., and Buchanan, R. L., and Miller, A. J. 2004. Potential for internalization, growth, and survival of *Salmonella* and *Escherichia coli* O157: H7 in oranges. J. Food Prot. 67:1578-1584.

El-Khawas, H., and Adachi, K. 1999. Identification and quantification of auxins in culture media of *Azospirillum* and *Klebsiella* and their effect on rice roots. Biol. Fertil. Soils 28:377-381.

Emsley, P., Charles, I. G., Fairweather, N. F., and Isaacs, N. W. 1996. Structure of *Bordetella pertussis* virulence factor P.69 pertactin. Nature 381:90-92.

Engelhard, M., Hurek, T., and Reinhold-Hurek B. 2000. Preferential occurrence of diazotrophic endophytes, *Azoarcus* spp., in wild rice species and land races of *Oryza sativa* in comparison with modern races. Environ. Microbiol. 2:131-141.

Espinosa-Urgel, M., and Ramos, J. L. 2004. Cell density-dependent gene contributes to efficient seed colonization by *Pseudomonas putida* KT2440. Appl. Environ. Microbiol. 70:5190-5198.

Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783-791.

Finkelstein, R. R. 1994. Mutations at two new *Arabidopsis* ABA response loci are similar to the *abi3* mutations. Plant J. 5:765-771.

Fisher, M. M. and Triplett, E. W. 1999. Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities. Appl. Environ. Microbiol. 65:4630-4636.

Folimonova, S. Y., Robertson, C. J., Garnsey, S. M., Gowda, S., and Dawson. W. O. 2009. Examination of the responses of different genotypes of citrus to Huanglongbing (citrus greening) under different conditions. Phytopathology. In press.

Fournet-Fayard, S., Joly, B., and Forestier, C. 1995. Transformation of wild type *Klebsiella pneumoniae* with plasmid DNA by electroporation. J. Microbiol. Methods 24:49-54.

Fouts, D. E., Tyler, H. L., DeBoy, R. T., Daugherty, S., Ren, Q., Badger, J. H., Durkin, A. S., Huot, H., Shrivastava, S., Kothari, S., Dodson, R. J., Mohamoud, Y., Khouri, H., Roesch, L. F. W., Krogfelt, K. A., Struve, C., Triplett, E. W., and Methé, B. A. 2008. Complete genome sequence of the N_2-fixing broad host range endophyte *Klebsiella pneumoniae* 342 and virulence predictions verified in mice. PLoS Genet. 4:e1000141.

Franz, E., van Diepeningen, A. D., de Vos, O. J., van Bruggen, A. H. C. 2005. Effects of cattle feeding regimen and soil management type on the fate of *Escherichia coli* O157: H7 and *Salmonella enterica* serovar typhimurium in manure, manure-amended soil, and lettuce. Appl. Environ. Microbiol. 71:6165-6174.

Franz, E., Visser, A. A., van Diepeningen, A. D., Klerks, M. M., Termorshuizen, A. J., and van Bruggen, A. H. C. 2007. Quantification of contamination of lettuce by GFP-expressing *Escherichia coli* O157: H7 and *Salmonella enterica* serovar Typhimurium. Food Microbiol. 24:106-112.

Fuentes-Ramirez, L. E., Jimenez-Salgado, T., Abarcao-Campo, I. R., and Caballero-Mellado, J. 1993. *Acetobacter diazotrophicus*, an indoleacetic acid producing bacterium isolated from sugarcane cultivars of Mexico. Plant Soil 154:145-150.

Fulchieri, M., Lucangeli, C., and Bottini, R. 1993. Inoculation with *Azospirillum lipoferum* affects growth and gibberellin status of corn seedling roots. Plant Cell Physiol. 34:1305-1309.

Gagliardi, J. V., and Karns, J. S. 2002. Persistence of *Escherichia coli* O157: H7 in soil and on plant roots. Environ. Microbiol. 4:89-96.

Garcia de Salamone, I. E., Hynes, R. K., and Nelson, L. M. 2001. Cytokinin production by plant growth promoting rhizobacteria and selected mutants. Can. J. Microbiol. 47:404-411.

Garnier, M. and Bové, J. M. 1983. Transmission of the organism associated with citrus greening disease from sweet orange to periwinkle by dodder. Phytopathology 73:1358-1363.

Garnier, M., Danel, N., and Bové, J. M. 1984. Aetiology of citrus greening disease. Annales De Microbiologie 135A:169-179.

Ghosh, S. K., Giannotti, J., and Louis, C. 1978. Multiplication intensive des prokaryotes associés aux maladies de type “greening” des agrumes dans les cellules criblées de cuscute. Annales de Phytopathologie 9:525-530.

Gillis, M., Kersters, K., Hoste, B., Janssens, D., Kroppenstedt, R. M., Stephan, M. P., Teixeira, K. R. S., Döbereiner, J., and De Ley, J. 1989. *Acetobacter diazotrophicus* sp. nov., a nitrogen-fixing acetic-acid bacterium associated with sugarcane. Int. J. Syst. Bacteriol. 39:361-364.

Girardin, H., Morris, C. E., Albagnac, C., Dreux, N., Glaux, C., and Nguyen-The, C. 2005. Behaviour of the pathogen surrogates *Listeria innocua* and *Clostridium sporogenes* during production of parsley in fields fertilized with contaminated amendments. FEMS Microbiol. Ecol. 54:287-295.

Glick, B. R. 2005. Modulation of plant ethylene levels by the bacterial enzyme ACC deaminase. FEMS Microbiol. Lett. 251:1-7.

Gordan, S. A., and Weber, R. P. 1951. Colorimetric estimation of indoleacetic acid. Plant Physiol. 26:192-195.

Govan, J. R. W., Hughes, J. E., and Vandamme, P. 1996. *Burkholderia cepacia*: medical, taxonomic and ecological issues. J. Med. Microbiol. 45:395-407.

Guo, X., Chen, J. R., Brackett, R. E., and Beuchat, L. R. 2001. Survival of Salmonellae on and in tomato plants from the time of inoculation at flowering and early stages of fruit development through fruit ripening. Appl. Environ. Microbiol. 67:4760-4764.

Gutierrez-Manero, F. J., Ramos-Solano, B., Probanza, A., Mehouachi, J., Tadeo, F. R., and Talon, M. 2001. The plant-growth-promoting rhizobacteria *Bacillus pumilus* and *Bacillus licheniformis* produce high amounts of physiologically active gibberellins. Physiol. Plant. 111:206-211.

Guzman, P. and Ecker, J. R. 1990. Exploiting the triple response of *Arabidopsis* to identify ethylene-related mutants. Plant Cell 2:513-523.

Gyaneshwar, P., James, E. K., Mathan, N., Reddy, P. M., Reinhold-Hurek, B., and Ladha, J. K. 2001. Endophytic colonization of rice by a diazotrophic strain of *Serratia marcescens*. J. Bacteriol. 183:2634-2645.

Hall, J. A., Peirson, D., Ghosh, S., and Glick, B. R. 1996. Root elongation in various agronomic crops by the plant growth promoting rhizobacterium *Pseudomonas putida* GR12-2. Isr. J. Plant Sci. 44:37-42.

Hardt, W. D., and Galan, J. E. 1997. A secreted *Salmonella* protein with homology to an avirulence determinant of plant pathogenic bacteria. Proc. Natl. Acad. Sci. USA 94:9887-9892.

Hassan, A. N., and Frank, J. F. 2004. Attachment of *Escherichia coli* O157: H7 grown in tryptic soy broth and nutrient broth to apple and lettuce surfaces as related to cell hydrophobicity, surface charge, and capsule production. Int. J. Food Microbiol. 96:103-109.

He, J. X., Baldini, R. L., Deziel, E., Saucier, M., Zhang, Q. H., Liberati, N. T., Lee, D., Urbach, J., Goodman, H. M., and Rahme, L. G. 2004. The broad host range pathogen *Pseudomonas aeruginosa* strain PA14 carries two pathogenicity islands harboring plant and animal virulence genes. Proc. Natl. Acad. Sci. USA 101:2530-2535.

Helling, R. B., Janes, B. K., Kimball, H., Tran, T., Bundesmann, M., Check, P., Phelan, D., and Miller, C. 2002. Toxic waste disposal in *Escherichia coli*. J. Bacteriol. 184:3699-3703.

Himanen, K., Boucheron, E., Vanneste, S., de Almeida Engler, J., Inze, D., and Beeckman, T. 2002. Auxin-mediated cell cycle activation during early lateral root initiation. Plant Cell 14:2339-2351.

Himanen, K., Vuylsteke, M., Vanneste, S., Vercruysse, S., Boucheron, E., Alard, P., Chriqui, D., van Montagu, M., Inze, D., and Beeckman, T. 2004. Transcript profiling of early lateral root initiation. Proc. Natl. Acad. Sci. USA 101:5146-5151.

Hobbie, L., and Estelle, M. 1995. The *axr4* auxin-resistant mutants of *Arabidopsis thaliana* define a gene important for root gravitropism and lateral root initiation. Plant J. 7:211-220.

Hoefsloot, G., Termorshuizen, A. J., Watt, D. A., and Cramer, M. D. 2005. Biological nitrogen fixation is not a major contributor to the nitrogen demand of a commercially grown South African sugarcane cultivar. Plant Soil 277:85-96.

Hontzeas, N., Saleh, S. S., and Glick, B. R. 2004. Changes in gene expression in canola roots induced by ACC-deaminase-containing plant-growth-promoting bacteria. Mol. Plant-Microbe. Interact. 17:865-8671.

Howard, M. B., and Hutcheson, S. W. 2003. Growth dynamics of *Salmonella enterica* strains on alfalfa sprouts and in waste seed irrigation water. Appl. Environ. Microbiol. 69:548-553.

Hrabak, E. M., and Willis, D. K. 1992. The *lemA* gene required for pathogenicity of *Pseudomonas syringae* pv. syringae on bean is a member of a family of two-component regulators. J. Bacteriol. 174:3011-3020.

Hrabak, E. M., and Willis D. K. 1993. Involvement of the *lemA* gene in production of syringomycin and protease by *Pseudomonas syringae* pv. *syringae*. Mol. Plant-Microbe Interact. 6:368-3675.

Ibekwe, A. M., Watt, P. M., Shouse, P. J., and Grieve, C. M. 2004. Fate of *Escherichia coli* O157: H7 in irrigation water on soils and plants as validated by culture method and real-time PCR. Can. J. Microbiol. 50:1007-1014.

Ibenyassine, K., AitMhand, R., Karamoko, Y., Cohen, N., and Ennaji, M. M. 2006. Use of repetitive DNA sequences to determine the persistence of enteropathogenic *Escherichia coli* in vegetables and in soil grown in fields treated with contaminated irrigation water. Lett. Appl. Microbiol. 43:528-533.

Ibenyassine, K., Mhand, R. A., Karamoko, Y., Anajjar, B., Chouibani, M., and Ennaji, M. M. 2007. Bacterial pathogens recovered from vegetables irrigated by wastewater in Morocco. J. Environ. Health 69:47-51.

Idriss, E. E., Makarewicz, O., Farouk, A., Rosner, K., Greiner, R., Bochow, H., Richter, T., and Borriss, R. 2002. Extracellular phytase activity of *Bacillus amyloliquefaciens* FZB45 contributes to its plant-growth-promoting effect. Microbiology 148:2097-2109.

Iglewski, B. H., and Kabat, D. 1975. NAD-dependent inhibition of protein synthesis by *Pseudomonas aeruginosa* toxin. Proc. Natl. Acad. Sci. USA 72:2284-2288.

Iniguez, A. L., Dong, Y. M., Carter, H. D., Ahmer, B. M. M., Stone, J. M., and Triplett, E. W. 2005. Regulation of enteric endophytic bacterial colonization by plant defenses. Mol. Plant-Microbe. Interact. 18:169-178.

Iniguez, A. L., Dong, Y., and Triplett, E. W. 2004. Nitrogen fixation in wheat provided by *Klebsiella pneumoniae* 342. Mol. Plant-Microbe Interact. 17:1078-1085.

Islam, M., Doyle, M. P., Phatak, S. C., Millner, P., and Jiang, X. P. 2005. Survival of *Escherichia coli* O157: H7 in soil and on carrots and onions grown in fields treated with contaminated manure composts or irrigation water. Food Microbiol. 22:63-70.

Islam, M., Morgan, J., Doyle, M. P., Phatak, S. C., Millner, P., and Jiang, X. P. 2004. Fate of *Salmonella enterica* serovar Typhimurium on carrots and radishes grown in fields treated with contaminated manure composts or irrigation water. Appl. Environ. Microbiol. 70:2497-2502.

Itoh, Y., Sugita-Konishi, Y., Kasuga, F., Iwaki, M., Hara-Kudo, Y., Saito, N., and Noguchi, Y. 1998. Enterohemorrhagic *Escherichia coli* O157: H7 present in radish sprouts. Appl. Environ. Microbiol. 64:1532-1535.

Jablasone, J., Warriner, K., and Griffiths, M. 2005. Interactions of *Escherichia coli* O157: H7, *Salmonella typhimurium* and *Listeria monocytogenes* plants cultivated in a gnotobiotic system. Int. J. Food Microbiol. 99:7-18.

Jacob-Dubuisson, F., Locht, C., and Antoine, R. 2001. Two-partner secretion in Gram-negative bacteria: a thrifty, specific pathway for large virulence proteins. Mol. Microbiol. 40:306-313.

Jagoueix, S., Bové, J. M., and Garnier, M. 1994. The phloem-limited bacterium of greening disease of citrus is a member of the α subdivision of the *Proteobacteria*. Int. J. Syst. Bacteriol. 44:379-386.

Jagoueix, S., Bové, J. M., and Garnier, M. 1997. Comparison of the 16S/23S ribosomal intergenic regions of "*Candidatus* Liberobacter asiaticum" and "*Candidatus* Liberobacter africanum," the two species associated with citrus Huanglongbing (greening) disease. Int. J. Syst. Bacteriol. 47:224-227.

Janes, M. E., Kim, K. S., and Johnson, M. G. 2005. Transmission electron microscopy study of enterohemorrhagic *Escherichia coli* O157: H7 in apple tissue. J. Food Prot. 68:216-224.

Jantama, K., Zhang, X., Moore, J. C., Shanmugam, K. T., Svoronos, S. A., and Ingram, L. O. 2008. Eliminating side products and increasing succinate yields in engineered strains of *Escherichia coli* C. Biotechnol. Bioeng. 101:881-893.

Jeter, C., and Matthysse, A. G. 2005. Characterization of the binding of diarrheagenic strains of *E. coli* to plant surfaces and the role of curli in the interaction of the bacteria with alfalfa sprouts. Mol. Plant-Microbe Interact. 18:1235-1242.

Jha, A. K., Bais, H. P., and Vivanco, J. M. 2005. *Enterococcus faecalis* mammalian virulence-related factors exhibit potent pathogenicity in the *Arabidopsis thaliana* plant model. Infect. Immun. 73:464-475.

Jimenez, B., Austin, A., Cloete, E., and Phasha, C. 2006. Using Ecosan sludge for crop production. Water Sci. Technol. 54:169-177.

Jimenez-Salgado, T., Fuentes-Ramirez, L. E., Tapia-Hernandez, A., Mascarua-Esparza, M. A., Martinez-Romero, E., and Caballero-Mellado, J. 1997. *Coffea arabica* L., a new host plant for *Acetobacter diazotrophicus*, and isolation of other nitrogen-fixing acetobacteria. Appl. Environ. Microbiol. 63:3676-3683.

Johannessen, G. S., Bengtsson, G. B., Heier, B. T., Bredholt, S., Wasteson, Y., and Rorvik, L. M. 2005. Potential uptake of *Escherichia coli* O157: H7 from organic manure into crisphead lettuce. Appl. Environ. Microbiol. 71:2221-2225.

Kanbe, K., Naganawa, H., Okamura, M., Sasaki, T., Hamada, M., Okami, Y., and Takeuchi, I. 1993. Amidenin, a new plant growth-regulating substance isolated from *Amycolatopsis* sp. Biosci. Biotechnol. Biochem. 57:1261-1263.

Karadeniz, A., Topcuoglu, S. F., and Inan, S. 2006. Auxin, gibberellin, cytokinin and abscisic acid production in some bacteria. World J. Microbiol. Biotechnol. 22:1061-1064.

Kenney, S. J., Anderson, G. L., Williams, P. L., Millner, P. D., and Beuchat, L. R. 2006. Migration of *Caenorhabditis elegans* to manure and manure compost and potential for transport of *Salmonella newport* to fruits and vegetables. Int. J. Food Microbiol. 106:61-68.

Kenney, S. J., Burnett, S. L., and Beuchat, L. R. 2001. Location of *Escherichia coli* O157: H7 on and in apples as affected by bruising, washing, and rubbing. J. Food Prot. 64:1328-1333.

Klerks, M. M., Franz, E., van Gent-Pelzer, M., Zijlstra, C., van Bruggen, A. H. C. 2007. Differential interaction of *Salmonella enterica* serovars with lettuce cultivars and plant-microbe factors influencing the colonization efficiency. ISME J. 1:620-631.

Knee, E. M., and Hangarter, R. P. 1996. Differential IAA dose response relations of the *axr1* and *axr2* mutants of *Arabidopsis*. Physiol. Plant. 98:320-324.

Koga, J., Syono, K., Ichikawa, T., and Adachi, T. 1994. Involvement of L-tryptophan aminotransferase in indole-3-acetic acid biosynthesis in *Enterobacter cloacae*. Biochim. Biophys. Acta 1209:241-247.

Koornneef, M., Elgersma, A., Hanhart, C. J., van Leonen-Martinet, E. P., van Rign, L., and Zeevaart, J. A. D. 1985. A gibberellin insensitive mutant of *Arabidopsis thaliana*. Physiol. Plant. 65:33-39.

Koornneef, M., Reuling, G., and Karssen, C. M. 1984. The isolation and characterization of abscisic acid-insensitive mutants of *Arabidopsis thaliana*. Physiol. Plant. 61:377-383.

Krause, A., Ramakumar, A., Bartels, D., Battistoni, F., Bekel, T., Boch, J., Böhm, M., Friedrich, F., Hurek, T., Krause, L., Linke, B., McHardy, A. C., Sarkar, A., Schneiker, S., Syed, A. A., Thauer, R., Vorhölter, F. J., Weidner, S., Pühler, A., Reinhold-Hurek, B., Kaiser, O., and Goesmann, A. 2006. Complete genome of the mutualistic, N_2-fixing grass endophyte *Azoarcus* sp strain BH72. Nat. Biotechnol. 24:1385-1391.

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

Kutter, S., Hartmann, A., and Schmid, M. 2006. Colonization of barley (*Hordeum vulgare*) with *Salmonella enterica* and *Listeria* spp. FEMS Microbiol. Ecol. 56:262-271.

Lacava, P. T., Araujo, W. L., and Azevedo, J. L. 2007. Evaluation of endophytic colonization of *Citrus sinensis* and *Catharanthus roseus* seedlings by endophytic bacteria. J. Microbiol. 45:11-14.

Lai, Y. C., Peng, H. L., and Chang, H. Y. 2001. Identification of genes induced in vivo during *Klebsiella pneumoniae* CG43 infection. Infect. Immun. 69:7140-7145.

Lambrecht, M., Okon, Y., Vande Broek, A., and Vanderleyden, J. 2000. Indole-3-acetic acid: a reciprocal signaling molecule in bacteria-plant interactions. Trends Microbiol. 8:298-300.

Landini, P. and Volkert, M. R. 1995. Transcriptional activation of the *Escherichia coli* adaptive response gene *aidB* is mediated by binding of methylated Ada protein. J. Biol. Chem. 270:8285-8289.

Latreille, P., Norton, S., Goldman, B. S., Henkhaus, J., Miller, N., Barbazuk, B., Bode, H. B., Darby, C., Du, Z. J., Forst, S., Gaudriault, S., Goodner, B., Goodrich-Blair, H., and Slater, S. 2007. Optical mapping as a routine tool for bacterial genome sequence finishing. BMC Genomics 8:321.

Laville, J., Voisard, C., Keel, C., Maurhofer, M., Defago, G., and Haas, D. 1992. Global control in *Pseudomonas fluorescens* mediating antibiotic synthesis and suppression of black root rot of tobacco. Proc. Natl. Acad. Sci. USA 89:1562-1566.

Lawlor, M. S., Hsu, J., Rick, P. D., and Miller, V. L. 2005. Identification of *Klebsiella pneumoniae* virulence determinants using an intranasal infection model. Mol. Microbiol. 58:1054-1073.

Lebuhn, M., Heulin, T., and Hartmann, A. 1997. Production of auxin and other indolic and phenolic compounds by *Paenibacillus polymyxa* strains isolated from different proximity to plant roots. FEMS Microbiol. Ecol. 22:325-334.

Lee, S., Flores-Encarnación, M., Contreras-Zentella, M., Garcia-Flores, L., Escamilla, J. E., and Kennedy, C. 2004. Indole-3-acetic acid biosynthesis is deficient in *Gluconacetobacter diazotrophicus* strains with mutations in cytochrome c biogenesis genes. J. Bacteriol. 186:5384-5391.

Li, J. P., Ovakim, D. H., Charles, T. C., and Glick, B. R. 2000. An ACC deaminase minus mutant of *Enterobacter cloacae* UW4 no longer promotes root elongation. Curr. Microbiol. 41:101-105.

Lim, A., Dimalanta, E. T., Potamousis, K. D., Yen, G., Apodoca, J., Tao, C. H., Lin, J. Y., Qi, R., Skiadas, J., Ramanathan, A., Perna, N. T., Plunkett, G., Burland, V., Mau, B., Hackett, J., Blattner, F. R., Anantharaman, T. S., Mishra, B., and Schwartz, D. C. 2001. Shotgun optical maps of the whole *Escherichia coli* O157 : H7 genome. Genome Res. 11:1584-1593.

Lin, H., Doddapaneni, H., Bai, X. J., Yao, J. Q., Zhao, X. L., and Civerolo, E. L. 2008. Acquisition of uncharacterized sequences from *Candidatus* Liberibacter, an unculturable bacterium, using an improved genomic walking method. Mol. Cell. Probes 22:30-37.

Lincoln, C., Britton, J. H., and Estelle, M. 1990. Growth and development of the *axr1* mutants of *Arabidopsis*. Plant Cell 2:1071-1080.

Loncarevic, S., Johannessen, G. S., and Rorvik, L. M. 2005. Bacteriological quality of organically grown leaf lettuce in Norway. Lett. Appl. Microbiol. 41:186-189.

Lopez-Bucio, J., Hernandez-Abreu, E., Sanchez-Calderon, L., Nieto-Jacobo, M. F., Simpson, J., and Herrera-Estrella, L. 2002. Phosphate availability alters architecture and cause changes in hormone sensitivity in the *Arabidopsis* root system. Plant Physiol. 129:244-256.

Loubens, I., Debarbieux, L., Bohin, A., Lacroix, J. M., and Bohin, J. P. 1993. Homology between a genetic locus (*mdoA*) involved in the osmoregulated biosynthesis of periplasmic glucans in *Escherichia coli* and a genetic locus (*hrpM*) controlling pathogenicity of *Pseudomonas syringae*. Mol. Microbiol. 10:329-340.

Lynch, M., Painter, J., Woodruff, R., and Braden, C. 2006. Surveillance for foodborne-disease outbreaks-United States, 1998-2002. MMWR Surveill. Summ. 55:1-34.

Ma, D., Cook, D. N., Alberti, M., Pon, N. G., Nikaido, H., and Hearst, J. E. 1995. Genes *acrA* and *acrB* encode a stress-induced efflux system in *Escherichia coli*. Mol. Microbiol. 16:45-55.

Madigan, M. T., and Martino, J. M. 2006. Brock Biology of Microorganisms. 11th ed. Pearson Prentice Hall, Upper Saddle River, NJ.

Mahillon, J., and Chandler, M. 1998. Insertion sequences. Microbiol. Mol. Biol. Rev. 62:725-774.

Malamy, J. E., and Benfey, P. N. 1997. Organization and cell differentiation in lateral roots of *Arabidopsis thaliana*. Development 124:33-44.

Malamy, J. E., and Ryan, K. S. 2001. Environmental regulation of lateral root initiation in *Arabidopsis*. Plant Physiol. 127:899-909.

Markova, Y. A., Romanenko, A. S., and Dukhanina, A. V. 2005. Isolation of bacteria of the family Enterobacteriaceae from plant tissues. Microbiology 74:575-578.

Matos, A., and Garland, J. L. 2005. Effects of community versus single strain inoculants on the biocontrol of *Salmonella* and microbial community dynamics in alfalfa sprouts. J. Food Prot. 68:40-48.

McClean, A. P. D. and Oberholzer, P. C. J. 1965. Citrus psylla, a vector of the greening disease of sweet orange. S. Afr. J. Agric. Sci. 8:298-298.

Melloul, A. A., Hassani, L., and Rafouk, L. 2001. *Salmonella* contamination of vegetables irrigated with untreated wastewater. World J. Microbiol. Biotechnol. 17:207-209.

Meng, X., Benson, K., Chada, K., Huff, E. J., and Schwartz, D. C. 1995. Optical mapping of bacteriophage-lambda clones using restriction endonucleases. Nat. Genet. 9:432-438.

Mikulasova, M., Vaverkova, S., Birosova, L., and Suchanova, M. 2005. Genotoxic effects of the hydroxycinnamic acid derivatives – caffeic, chlorogenic and cichoric acids. Biologia 60:275-279.

Molina, M. A., Ramos, J. L., and Espinosa-Urgel, M. 2006. A two-partner secretion system is involved in seed ad root colonization and iron uptake by *Pseudomonas putida* KT2400. Environ. Microbiol. 8:639-647.

Monier, J. M., and Lindow, S. E. 2005. Aggregates of resident bacteria facilitate survival of immigrant bacteria on leaf surfaces. Microb. Ecol. 49:343-352.

Muirhead, R. W., Collins, R. P., and Bremer, P. J. 2006. Interaction of *Escherichia coli* and soil particles in runoff. Appl. Environ. Microbiol. 72:3406-3411.

Nakabeppu, Y., and Sekiguchi, M. 1986. Regulatory mechanisms for induction of synthesis of repair enzymes in response to alkylating agents: Ada protein acts as a transcriptional regulator. Proc. Nat. Acad. Sci. USA 83:6297-6301.

Nieto, K. F., and Frankenberger, W. T. Jr. 1989. Biosynthesis of cytokinins by *Azotobacter chroococcum*. Soil Biol. Biochem. 21:967-972.

Nodzon, L. A., Xu, W. H., Wang, Y., Pi, L. Y., Chakrabarty, P. K., and Song, W. Y. 2004. The ubiquitin ligase XBAT32 regulates lateral root development in *Arabidopsis*. Plant J. 40:996-1006.

Ohman, D. E., Sadoff, J. C., and Iglewski, B. H. 1980. Toxin A-deficient mutants of *Pseudomonas aeruginosa* PA103: isolation and characterization. Infect. Immun. 28:899-908.

Opelt, K., Berg, C., and Berg, G. 2007. The bryophyte genus *Sphagnum* is a reservoir for powerful and extraordinary antagonists and potentially facultative human pathogens. FEMS Microbiol. Ecol. 61:38-53.

Ostroff, R. M., and Vasil, M. L. 1987. Identification of a new phospholipase C activity by analysis of an insertional mutation in the hemolytic phospholipase C structural gene of *Pseudomonas aeruginosa*. J. Bacteriol. 169:4597-4601.

Overbeek, R., Begley, T., Butler, R. M., Choudhuri, J. V., Chuang, H. Y., Cohoon, M., de Crécy-Lagard, V., Diaz, N., Disz, T., Edwards, R., Fonstein, M., Frank, E. D., Gerdes, S., Glass, E. M., Goesmann, A., Hanson, A., Iwata-Reuyl, D., Jensen, R., Jamshidi, N., Krause, L., Kubal, M., Larsen, N., Linke, B., McHardy, A. C., Meyer, F., Neuweger, H., Olsen, G., Olson, R., Osterman, A., Portnoy, V., Pusch, G. D., Rodionov, D. A., Rückert, C., Steiner, J., Stevens, R., Thiele, I., Vassieva, O., Ye, Y., Zagnitko, O., and Vonstein, V. 2005. The subsystem approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic Acids Res. 33:5691-5702.

Park, S. J., and Gunsalus, R. P. 1995. Oxygen, iron, carbon, and superoxide control of the fumerase *fumA* and *fumC* genes of *Escherichia coli:* Role of the *acrA, fnr,* and *soxR* gene products. J. Bacteriol. 177:6255-6262.

Patten, C. L., and Glick, B. R. 1996. Bacterial biosynthesis of indole-3-acetic acid. Can. J. Microbiol. 42:207-220.

Patten, C. L., and Glick, B. R. 2002. Role of *Pseudomonas putida* indoleacetic acid in development of the host plant root system. Appl. Environ. Microbiol. 68:3795-3801.

Paula, M. A., Reis, V. M., and Döbereiner, J. 1991. Interactions of *Glomus clarum* with *Acetobacter diazotrophicus* in infection of sweet potato (*Ipomoea batatas*), sugarcane (*Saccharum* spp.), and sweet sorghum (*Sorghum vulgare*). Biol. Fert. Soils 11:111-115.

Pedrosa, F., and Consortium, G. 2005. A report on the genome of *Herbaspirillum seropedicae* strain Z78. Pages 111-114 in: Biological Nitrogen Nixation, Sustainable Agriculture and the Environment; Proceedings of the 14th International Nitrogen Fixation Congress. Beijing, People's Republic of China. October 27 - November 01, 2004. Springer, Dordrecht, Netherlands.

Peek, J.A., and Taylor, R. K. 1992. Characterization of a periplasmic thiol:disulfide interchange protein required for the functional maturation of secreted virulence factors of *Vibrio cholerae*. Proc. Natl. Acad. Sci. USA 89:6210-6214.

Persello-Cartieaux, F., David, P., Sarrobert, C., Thibaud, M. C., Achouak, W., Robglia, C., and Nussaume, L. 2001. Utilization of mutants to analyze the interaction between *Arabidopsis thaliana* and its naturally root-associated *Pseudomonas*. Planta 212:190-198.

Pieterse, C. M. J., van Wees, S. C. M., van Pelt, J. A., Knoester, M., Laan, R., Gerrits, N., Weisbeek, P. J., and van Loon, L. C. 1998. A novel signaling pathway controlling induced systemic resistance in Arabidopsis. Plant Cell 10:1571-1580.

Plotnikova, J. M., Rahme, L. G., and Ausubel, F. M. 2000. Pathogenesis of the human opportunistic pathogen *Pseudomonas aeruginosa* PA14 in *Arabidopsis*. Plant Physiol. 124:1766-1774.

Polz, M. F. and Cavanaugh, C. M. 1998. Bias in template-to-product ratios in multitemplate PCR. Appl. Environ. Microbiol. 64:3724-3730.

Rahme, L. G., Ausubel, F. M., Cao, H., Drenkard, E., Goumnerov, B. C., Lau, G. W., Mahajan-Miklos, S., Plotnikova, J., Tan, M. W., Tsongalis, J., Walendziewicz, C. L., and Tompkins, R. G. 2000. Plants and animals share functionally common bacterial virulence factors. Proc. Natl. Acad. Sci. USA 97:8815-8821.

Rahme, L. G., Stevens, E. J., Wolfort, S. F., Shao, J., Tompkins, R. G., and Ausubel, F. M. 1995. Common virulence factors for bacterial pathogenicity in plants and animals. Science 268:1899-1902.

Ramirez-Chavez, E., Lopez-Bucio, J., Herrera-Estrella, L., and Molina-Torres, J. 2004. Alkamides isolated from plants promote growth and alter root development in *Arabidopsis*. Plant Physiol. 134:1058-1068.

Ranjard, L., Poly, F., Lata, J. C., Mougel, C., Thioulouse, J., and Nazaret, S. 2001. Characterization of bacterial and fungal soil communities by automated ribosomal intergenic spacer analysis fingerprints: Biological and methodological variability. Appl. Environ. Microbiol. 67:4479-4487.

Raskin, I. 1992. Salicylate, A New Plant Hormone. Plant Physiol. 99:799-803.

Ray, S. K., Rajeshwari, R., Sharma, Y., and Sonti, R. V. 2002. A high-molecular-weight outer membrane protein of *Xanthomonas oryzae* pv. *oryzae* exhibits similarity to nonfimbrial adhesins of animal pathogenic bacteria and is required for optimum virulence. Mol. Microbiol. 46:637-647.

Regué, M., Enfedaque, J., Camprubi, S., and Tomás, J. M. 1992. The O-antigen lipopolysaccharide is the major barrier to plasmid DNA uptake by *Klebsiella pneumoniae* during transformation by electroporation and osmotic shock. J. Microbiol. Methods 15:129-134.

Reinhold-Hurek, B., Maes, T., Gemmer, S., Van Montagu, M., and Hurek, T. 2006. An endoglucanase is involved in infection of rice roots by the not-cellulose-metabolizing endophyte *Azoarcus* sp strain BH72. Mol. Plant-Microbe Interact. 19:181-188.

Reiter, B., Bürgmann, H., Burg, K., and Sessitsch, A. 2003. Endophytic *nifH* gene diversity in African sweet potato. Can. J. Microbiol. 49:549-555.

Reslewic, S., Zhou, S., Place, M., Zhang, Y. P., Briska, A., Goldstein, S., Churas, C., Runnheim, R., Forrest, D., Lim, A., Lapidus, A., Han, C. S., Roberts, G. P., and Schwartz, D. C. 2005. Whole-genome shotgun optical mapping of *Rhodospirillum rubrum*. Appl. Environ. Microbiol. 71:5511-5522.

Rich, J. J., Kinscherf, T. G., Kitten, T., and Willis, D. K. 1994. Genetic evidence that the *gacA* gene encodes the cognate response regulator for the *lemA* sensor in *Pseudomonas syringae*. J. Bacteriol. 176:7468-7475.

Riggs, P. J., Chelius, M. K., Iniguez, A. L., Kaeppler, S. M., and Triplett, E. W. 2001. Enhanced maize productivity by inoculation with Diazotrophic bacteria. Aust. J. Plant Physiol. 28:829-836.

Riggs, P. J., Moritz, R. L., Chelius, M. K., Dong, Y., Iniguez, A. L., Kaeppler, S. M., Casler, M. D., and Triplett, E. W. 2002. Isolation and characterization of diazotrophic endophytes from grasses and their effects on plant growth. Pages 263-267 in: Nitrogen fixation: Global perspectives: Proceedings of the 13th International Congress on Nitrogen Fixation. Hamilton, Ontario, Canada. July 2-7, 2001. CABI Publishing, New York, NY.

Robertson, L. J., Johannessen, G. S., Gjerde, B. K., and Loncarevic, S. 2002. Microbiological analysis of seed sprouts in Norway. Int. J. Food Microbiol. 75:119-126.

Roesch, L. F., Fulthorpe, R. R., Riva, A., Casella, G., Hadwin, A. K. M., Kent, A. D., Daroub, S. H., Camargo, F. A. O., Farmerie, W. G., and Triplett, E. W. 2007. Pyrosequencing enumerates and contrasts soil microbial diversity. ISME J. 1:283-290.

Rojas, C. M., Ham, J. H., Deng, W. L., Doyle, J. J., and Collmer, A. 2002. HecA, a member of a class of adhesins produced by diverse pathogenic bacteria, contributes to the attachment, aggregation, epidermal cell killing, and virulence phenotypes of *Erwinia chrysanthemi* EC16 on *Nicotiana clevelandii* seedlings. Proc. Natl. Acad. Sci. USA 99:13142-13147.

Rosenblueth, M., Martinez, L., Silva, J., and Martinez-Romero, E. 2004. *Klebsiella variicola*, a novel species with clinical and plant-associated isolates. Syst. Appl. Microbiol. 27:27-35.

Rosenblueth, M., and Martinez-Romero, E. 2006. Bacterial endophytes and their interactions with hosts. Mol. Plant-Microbe Interact. 19:827-837.

Rupp, W. D., Sancar, A., and Sancar, G. B. 1982. Properties and regulation of the UvrABC endonuclease. Biochimie 64:595-598.

Ryu, C. M., Fang, M. A., Hu, C. H., Reddy, M. S., Wei, H. X., Pare, P. W., and Kloepper, J. W. 2003. Bacterial volatiles promote growth in *Arabidopsis*. Proc. Natl. Acad. Sci. USA. 100:4927-4932.

Sagaram, U. S., DeAngelis, K. M., Trivedi, P., Andersen, G. L., Lu, S. E., and Wang, N. 2009. Bacterial diversity analysis of Huanglongbing pathogen-infected citrus, using PhyloChip arrays and 16S rRNA gene clone library sequencing. Appl. Environ. Microbiol. 75:1566-1574.

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

Sevilla, M., Burris, R. H., Gunapala, N., and Kennedy, C. 2001. Comparison of benefit to sugarcane plant growth and $^{15}N_2$ incorporation following inoculation of sterile plants with *Acetobacter diazotrophicus* wild-type and Nif- mutant strains. Mol. Plant-Microbe. Interact. 14:358-366.

Shevchik, V. E., Bortoli-German, I., Robert-Baudouy, J., Robinet, S., Barras, F., and Condemine, G. 1995. Differential effect of *dsbA* and *dsbC* mutations on extracellular enzyme secretion in *Erwinia chrysanthemi*. Mol. Microbiol. 16:745-753.

Solomon, E. B., and Matthews, K. R. 2005. Use of fluorescent microspheres as a tool to investigate bacterial interactions with growing plants. J. Food Prot. 68:870-873.

Solomon, E. B., and Matthews, K. R. 2006. Interaction of live and dead *Escherichia coli* O157: H7 and fluorescent microspheres with lettuce tissue suggests bacterial processes do not mediate adherence. Lett. Appl. Microbiol. 42:88-93.

Solomon, E. B., Yaron, S., and Matthews, K. R. 2002. Transmission of *Escherichia coli* O157: H7 from contaminated manure and irrigation water to lettuce plant tissue and its subsequent internalization. Appl. Environ. Microbiol. 68:397-400.

Song, I., Stine, S. W., Choi, C. Y., and Gerba, C. P. 2006. Comparison of crop contamination by microorganisms during subsurface drip and furrow irrigation. J. Environ. Eng. ASCE 132:1243-1248.

Staswick, P. E., Su, W., and Howell, S. H. 1992. Methyl jasmonate inhibition of root growth and induction of a leaf protein are decreased in an *Arabidopsis thaliana* mutant. Proc. Natl. Acad. Sci. U.S.A. 89:6837-6840.

Struve, C., Forestier, C., and Krogfelt, K. A. 2003. Application of a novel multi-screening signature-tagged mutagenesis assay for identification of *Klebsiella pneumoniae* genes essential in colonization and infection. Microbiology 149:167-176.

Surette, M. A., Sturz, A. V., Lada, R. R., and Nowak, J. 2003. Bacterial endophytes in processing carrots (*Daucus carota* L. var. *sativus*): their localization, population density, biodiversity and their effects on plant growth. Plant Soil 253:381-390.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. 2007. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24:1596-1599.

Tapia-Hernandez, A., Bustillos-Cristales, M. R., Jimenez-Salgado, T., Caballero-Mellado, J., and Fuentes-Ramirez, L. E. 2000. Natural endophytic occurrence of *Acetobacter diazotrophicus* in pineapple plants. Microb. Ecol. 39:49-55.

Tatineni, S., Sagaram, U. S., Gowda, S., Robertson, C. J., Dawson, W. O., Iwanami, T., and Wang, N. 2008. In planta distribution of '*Candidatus* Liberibacter asiaticus' as revealed by polymerase chain reaction (PCR) and real-time PCR. Phytopathology 98:592-599.

Teixeira, D. C., Saillard, C., Eveillard, S., Danet, J. L., da Costa, P. I., Ayres, A. J., and Bové, J. 2005. '*Candidatus* Liberibacter americanus', associated with citrus huanglongbing (greening disease) in Sao Paulo State, Brazil. Int. J. Syst. Evol. Microbiol. 55:1857-1862.

Teixeira, D. C., Wulff, E. C., Martins, E. W., Kitajima R., Bassanezi, R., Ayres, A. J., Eveillard, S., Saillard, C., and Bové, J. M. 2008. A phytoplasma closely related to pigeon pea witches'-broom phytoplasma (16Sr IX) is associated with citrus huanglongbing symptoms in the state of São Paulo, Brazil. Phytopathology 98:977-984.

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

Tyler, H. L., Roesch, L. F. W., Gowda, S., Dawson, W. O., and Triplett, E. W. 2009. Confirmation of the sequence of '*Candidatus* Liberibacter asiaticus' and assessment of microbial diversity in Huanglongbing-infected citrus phloem using a metagenomic approach. Mol. Plant-Microbe Interact. In press.

Tyler, H. L., and Triplett, E. W. 2008. Plants as a habitat for beneficial and/or human pathogenic bacteria. Annu. Rev. Phytopathol. 46:53-73.

Timmusk, S., Nicander, B., Granhall, U., and Tillberg, E. 1999. Cytokinin production by *Paenibacillus polymyxa*. Soil Biol. Biochem. 31:1847-1852.

Torres, A. G., Jeter, C., Langley, W., and Matthysse, A. G. 2005. Differential binding of *Escherichia coli* O157: H7 to alfalfa, human epithelial cells, and plastic is mediated by a variety of surface structures. Appl. Environ. Microbiol. 71:8008-8015.

Ueda, Y., Yumoto, N., Tokushige, M., Fukui, K., and Ohya-Nishiguchi, H. 1991. Purification and characterization of two types of fumarase from *Escherichia coli*. J. Biochem. 109:728-733.

Vanneste, S., De Rybel, B., Beemster, G. T. S., Ljung, K., De Smet, I., Van Insterdael, G., Naudts, M., Iida, R., Gruissem, W., Tasaka, M., Inze, D., Fukaki, H., and Beeckman, T. 2005. Cell cycle progression in the pericycle is not sufficient for SOLITARY ROOT/IAA14-mediated lateral root initiation in *Arabidopsis thaliana*. Plant Cell 17:3035-3050.

Vasilieva, S. V., and Moschkovskaya, E. J. 2005. Quasi-adaptive response to alkylating agents in *Escherichia coli*: A new phenomenon. Russ. J. Genet. 41:484-489.

Wachtel, M. R., Whitehand, L. C., and Mandrell, R. E. 2002. Association of *Escherichia coli* O157: H7 with preharvest leaf lettuce upon exposure to contaminated irrigation water. J. Food Prot. 65:18-25.

Wang, C. X., Knill, E., Glick, B. R., and Defago, G. 2000. Effect of transferring 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase genes into *Pseudomonas fluorescens* strain CHA0 and its *gacA* derivative CHA96 on their growth-promoting and disease-suppressive capacities. Can. J. Microbiol. 46:898-907.

Wang, E. T., Tan, Z. Y., Guo, X. W., Rodriguez-Duran, R., Boll, G., and Martinez-Romero, E. 2006. Diverse endophytic bacteria isolated from a leguminous tree *Conzattia multiflora* grown in Mexico. Arch. Microbiol. 186:251-259.

Warriner, K., Spaniolas, S., Dickinson, M., Wright, C., and Waites, W. M. 2003. Internalization of bioluminescent *Escherichia coli* and *Salmonella* Montevideo in growing bean sprouts. J. Appl. Microbiol. 95:719-727.

Watarai, M., Tobe, T., Yoshikawa, M., and Sasakawa, C. 1995. Disulfide oxidoreductase activity of *Shigella flexneri* is required for release of Ipa proteins and invasion of epithelial cells. Proc. Natl. Acad. Sci. USA 92:4927-4931.

Weinstein, M., Roberts, R. C., and Helinski, D. R. 1992. A region of the broad-host range plasmid RK2 causes stable in planta inheritance of plasmids in *Rhizobium meliloti* cells isolated from alfalfa root nodules. J. Bacteriol. 174:7486-7489.

Williams, A. P., Avery, L. M., Killham, K., and Jones, D. L. 2007. Survival of *Escherichia coli* O157: H7 in the rhizosphere of maize grown in waste-amended soil. J. Appl. Microbiol. 102:319-326.

Williamson, L. C., Ribrioux, S. P. C. P., Fitter, A. H., and Ottoline-Leyser, H. M. 2001. Phosphate availability regulates root system architecture in *Arabidopsis*. Plant Physiol. 126:875-882.

Wilmoth, J. C., Wang, S., Tiwari, S. B., Joshi, A. D., Hagen, G., Guilfoyle, T. J., Alonso, J. M., Ecker, J. R., and Reed, J. W. 2005. NPH4/ARF7 and ARF19 promote leaf expansion and auxin-induced lateral root formation. Plant J. 43:118-130.

Wilson, A. K., Pickett, F. B., Turner, J. C., and Estelle, M. 1990. A dominant mutation in *Arabidopsis* confers resistance to auxin, ethylene, and abscisic acid. Molecular and General Genetics. 222:377-382.

Wood, B. E., Yomano, L. P., York, S. W., and Ingram, L. O. 2005. Development of industrial-medium-required elimination of the 2,3-butanediol fermentation pathway to maintain ethanol yield in an ethanologenic strain of *Klebsiella oxytoca*. Biotechnol. Prog. 21:1366-1372.

Wu, C. W., Schramm, T. M., Zhou, S., Schwartz, D. C., and Talaat, A. M. 2009. Optical mapping of the *Mycobacterium avium* subspecies *paratuberculosis* genome. BMC Genomics 10:25.

Xie, Q., Frugis, G., Cogan, D., and Chua, N. H. 2000. *Arabidopsis* NAC1 transduces auxin signal downstream of TIR1 to promote lateral root development. Genes Dev. 14:3024-3036.

Xie, Q., Guo, H. S., Dallman, G., Fang, S., Weissman, A. M., and Chua, N. H. 2002. SINAT5 promotes ubiquitin-related degradation of NAC1 to attenuate auxin signals. Nature 419:167-170.

Yorgey, P., Rahme, L. G., Tan, M. W., and Ausubel, F. M. 2001. The roles of *mucD* and alginate in the virulence of *Pseudomonas aeruginosa* in plants, nematodes and mice. Mol. Microbiol. 41:1063-1076.

Zelenev, V. V., van Bruggen, A. H. C., and Semenov, A. M. 2005. Short-term wavelike dynamics of bacterial populations in response to nutrient input from fresh plant residues. Microb. Ecol. 49:83-93.

Zhang, H., Jennings, A., Barlow, P. W., and Forde, B. G. 1999. Dual pathways for regulation of root branching by nitrate. Proc. Natl. Acad. Sci. USA. 96:6529-6534.

Zhou, S., Bechner, M. C., Place, M., Churas, C. P., Pape, L., Leong, S. A., Runnheim, R., Forrest, D. K., Goldstein, S., Livny, M., and Schwartz, D. C. 2007. Validation of rice genome sequence by optical mapping. BMC Genomics 8:278.

Zhou, S., Deng, W., Anantharaman, T. S., Lim, A., Dimalanta, E. T., Wang, J., Wu, T., Chunhong, T., Creighton, R., Kile, A., Kvikstad, E., Bechner, M., Yen, G., Garic-Stankovic, A., Severin, J., Forrest, D., Runnheim, R., Churas, C., Lamers, C., Perna, N. T., Burland, V., Blattner, F. R., Mishra, B., and Schwartz, D. C. 2002. A whole-genome shotgun optical map of *Yersinia pestis* strain KIM. Appl. Environ. Microbiol. 68:6321-6331.

Zhou, S., Kile, A., Bechner, M., Place, M., Kvikstad, E., Deng, W., Wei, J., Severin, J., Runnheim, R., Churas, C., Forrest, D., Dimalanta, E. T., Lamers, C., Burland, V., Blattner, F. R., and Schwartz, D. C. 2004a. Single-molecule approach to bacterial genomic comparisons via optical mapping. J. Bacteriol. 186:7773-7782.

Zhou, S., Kile, A., Kvikstad, E., Bechner, M., Severin, J., Forrest, D., Runnheim, R., Churas, C., Anantharaman, T. S., Myler, P., Vogt, C., Ivens, A., Stuart, K., and Schwartz, D. C. 2004b. Shotgun optical mapping of the entire *Leishmania major* Friedlin genome. Mol. Biochem. Parasitol. 138:97-106.

BIOGRAPHICAL SKETCH

Heather Tyler was born and raised in Lake Wales, Florida. She graduated from Lake Wales High School as co-salutatorian in the spring of 1999 and went on to attend Florida Southern College that fall, where she majored in biology and minored in chemistry. While attending Florida Southern, Heather became a member of the Phi Eta Sigma and Beta Beta Beta honor societies. She also participated in the school's chapter of Habitat for Humanity, serving as secretary her senior year. Heather graduated summa cum laude from Florida Southern College in the spring of 2003. That fall, she went on to pursue graduate studies in microbiology at the University of Florida, where her research focused on plant-associated bacteria and their roles in altering and enhancing plant growth. The goal of her work in graduate school has been to gain knowledge that will ultimately aid in increasing production of food crops and reducing the dependence of these plants on externally applied fertilizers, thereby reducing the environmental impact of agricultural land.