Investigation of *Pantoea stewartii* Quorum-Sensing Controlled Regulators and Genes Important for Infection of Corn

An Duy Duong

Dissertation submitted to the faculty of Virginia Tech in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In

Biological Sciences

_____	_____
Ann M. Stevens, PhD; Chair	Roderick V. Jensen, PhD
_____	_____
Stephen B. Melville, PhD	Florian D. Schubot, PhD

November 29th, 2017

Blacksburg, VA, USA

Keywords: quorum sensing, *Pantoea stewartii*, Stewart's wilt, RcsA, LrhA, regulon, transcription regulation, RNA-Seq, Tn-Seq, whole genome sequencing, phytopathogen

Investigation of *Pantoea stewartii* Quorum-Sensing Controlled Regulators and Genes Important for Infection of Corn

An Duy Duong

Dissertation submitted to the faculty of Virginia Tech in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In

Biological Sciences

Ann M. Stevens, PhD; Chair

Roderick V. Jensen, PhD

Stephen B. Melville, PhD

Florian D. Schubot, PhD

November 29th, 2017

Blacksburg, VA, USA

Keywords: quorum sensing, *Pantoea stewartii*, Stewart's wilt, RcsA, LrhA, regulon, transcription regulation, RNA-Seq, Tn-Seq, whole genome sequencing, phytopathogen

Investigation of *Pantoea stewartii* Quorum-Sensing Controlled Regulators and Genes Important for Infection of Corn

An Duy Duong

ABSTRACT

Bacteria interact with their eukaryotic hosts using a variety of mechanisms that range from being beneficial to detrimental. This dissertation focuses on *Pantoea stewartii* subspecies *stewartii* (*P. stewartii*), an endosymbiont in the corn flea beetle gut that causes Stewart's wilt disease in corn. Gaining insights into the interactions occurring between this bacterial pathogen and its plant host may lead to informed intervention strategies. This phytopathogen uses quorum sensing (QS) to coordinate cell density-dependent gene expression and successfully colonize corn leading to wilt disease. Prior to the research presented in this dissertation, the QS master regulator EsaR was shown to regulate two major virulence factors of *P. stewartii*, capsule production and surface motility. However, the function and integration of EsaR downstream targets in *P. stewartii* were still largely undefined. Moreover, only a draft genome of a reference strain of *P. stewartii* was publicly available for researchers, limiting bioinformatics and genome-scale genetic approaches with the organism. The work described in this dissertation has now addressed these important issues.

The function of two EsaR direct targets, LrhA and RcsA, was explored (Chapter Two) and the existence of integration in the regulation between them was discovered (Chapters Two and Four). RcsA and LrhA are transcription factors controlling capsule production and surface motility in *P. stewartii*, respectively. In Chapter Two, the RcsA and LrhA regulons were investigated using RNA-Seq. This led to the discovery of a potential regulatory interaction

between them that was confirmed by qRT-PCR and transcriptional gene fusion assays. The involvement of LrhA in surface motility and virulence was also established in this project. A direct interaction between LrhA and promoter of *rcsA* was defined in Chapter Four. Additional direct regulatory targets of LrhA were also identified.

A project to generate a complete assembly of the *P. stewartii* genome (Chapter Three) enabled more thorough genome-wide analysis and revealed the existence of a previous unknown 66-kb region in the *P. stewartii* genome believed to contain genes important for motility and virulence. In addition, completion of the genome sequence permitted genes for two distinctive Type III secretion systems, used for interactions with corn or the corn flea beetle, to be placed on two mega-plasmids. Furthermore, the complete genome sequence facilitated a Tn-Seq approach (Chapter Five). Tn-Seq is a potent tool used to identify bacterial genes required for certain environmental test conditions. This project is a pioneering utilization of a Tn-Seq analysis *in planta* to investigate genes important for colonization and survival of *P. stewartii* within its corn host. It was discovered that OmpC and Lon are important to *in planta* growth and OmpA plays a role in plant virulence.

In conclusion, these studies have broadened our understanding about the role of the QS regulon and other genes important for the pathogenesis of this phytopathogen. This knowledge may now be applied toward the development of future disease intervention strategies against *P. stewartii* and other wilt-disease causing plant pathogens.

ACKNOWLEDGEMENTS

First, I would like to thank the Graduate School and the Biological Sciences Department at Virginia Tech for providing me the opportunity to study my PhD with state-of-the-art technology.

I would like to express my deepest appreciation to my mentor, Dr. Ann M. Stevens, who has fully supported me through this journey of scientific training for my PhD degree in "microbiology and molecular biology" at the Department of Biological Sciences at Virginia Tech. She has been a great source of tutelage over the period of my development to become more independent as a scientist, as well as an educator. I believe that Dr. Stevens has trained me to be better in both aspects of the scientist that I want to become, strong in both research and teaching.

Next, I would like to thank Dr. Roderick V. Jensen who is not only a member in my scientific advisory committee but also a co-mentor who guided me during this period, especially with bioinformatics. He has kindled my interest in bioinformatics, which is a topic that I used to avoid. The more discussion we had, the more motivation I gained to keep improving my bioinformatic skills. I am now confident in using a variety of bioinformatic tools in my experimental research thanks to his guidance and discussion.

I am grateful to Dr. Stephen B. Melville and Dr. Florian D. Schubot, members of my scientific advisory committee. They have been extremely supportive during and outside each committee meeting, every semester. Their questions and advice were interesting and meaningful to me as I kept moving forward. I also appreciate the support from Dr. David Popham for the multiple ways he assisted me during my time working in the Life Science 1 building.

Additionally, I greatly appreciated working with members in Dr. Stevens' laboratory. Dr. Revathy Ramachandran, a senior PhD student at the time I joined the group, was a great source of support at the beginning of my first year in the laboratory. Dr. Alison Kernell Burke was the next senior lab-mate who helped me to become more independent. She and I shared a co-first authorship on the publication presented as Chapter Two in this dissertation. Holly Packard was a junior lab-mate who I have spent the last couple years with during conducting experiments. Ian

Hines joined the Stevens lab as a new graduate student and he is a good fit into the lab's atmosphere. I wish them both success in their career of choice. I had been also very pleased to work with and mentor the undergraduate students in our laboratory during my training here.

Moreover, I appreciate the Life Science I and Derring Hall communities, especially people working on the second and the fifth floors. Fellow graduate students including Manisha Shrestha, Cameron Sayer, Jordan Mancl, Benjamin Webb, Will Hendrick, Gary Camper, Andreas Sukmana, Bidisha Barat, Angie Estrada and others with whom I have had time to interact with, have been nice and treated me like a friend. I wish them all well with their current projects and their future/current careers.

In addition, it will not be completed if the department's teaching laboratories team was not mentioned here. I have had a very educational teaching experience at Virginia Tech. With that, I would like to thank Katrina Lasley, Sarah Nikraftar, Kevin Laoh, and other staff members as well as the other teaching assistants I worked with such as Elizabeth Kowalski, Rebecca Price, and Hualan Liu, among others. They have been very supportive and helped fulfill my teaching assistant position with great experiences. I am also thankful to all my students in the General Microbiology and Pathogenic Bacteria lab courses over the years.

Last but not least, I would not be able to meet and work with all those wonderful people, let alone to finish all the projects presented in this dissertation, without the support from my family: my parents, sisters, brother, niece, and especially my husband. I love you all. Their love and support are my best motivation to go far and beyond boundaries throughout my life.

ATTRIBUTIONS

I disclose that all experimental work reported herein, except for the statements below, were performed by myself. In addition, several colleagues contributed to the writing, data collection and analysis described in Chapters Two-Five as briefly reported here.

Ann M. Stevens, PhD, Professor in the Biological Sciences Department at Virginia Tech, was the research adviser for all projects represented in this dissertation. She conceived and designed experiments, contributed reagents/material/analysis tools, analyzed data, and assisted with writing for all Chapters.

Roderick V. Jensen, PhD, Professor in the Biological Sciences Department at Virginia Tech, was the co-research-adviser for all projects and the main adviser for project detailed in Chapter Three. He conceived and designed experiments, contributed reagents/material/analysis tools, and analyzed data for work done in Chapters Two, Three and Five, and assisted with writing for all Chapters.

Alison Kernell Burke, PhD, Instructional PT Temporary Adjunct in the Biological Sciences Department at Virginia Tech, was co-first author in the publication of the project described in Chapter Two. She conceived and designed experiments, analyzed data, and assisted with writing for the research in Chapter Two of this dissertation.

TABLE OF CONTENTS

LIST OF TABLES

Page number

Chapter Two

Chapter Three

Chapter Four

Chapter Five

Appendix A

Appendix B

Appendix C

LIST OF FIGURES

Page number

Appendix B

Appendix C

LIST OF ABBREVIATIONS

AHL	N-acyl homoserine lactone
Ap	Ampicillin
cDNA	Complementary DNA
CFU	Colony forming unit
Cm	Chloramphenicol
ddNTP	Dideoxynucleotide
EMSA	Electrophoretic mobility shift assays
EPS	Exopolysaccharide
GFP	Green fluorescent protein
HCD	High cell density
Hrp	Hypersensitive response and pathogenicity
HSL	Homoserine lactone
Kn	Kanamycin
LB	Luria-Bertani (medium)
LCD	Low cell density
MP	Mate-pair (library preparation for sequencing)
Nal	Nalidixic acid
NGS	Next generation sequencing
OD	Optical density
PBS	Phosphate buffered saline
PE	Paired-end (sequencing)
PTS	Phosphoenolpyruvate phosphotransferase system
qRT-PCR	Quantitative reverse transcription PCR
QS	Quorum sensing
RCI	Relative competition index
RM	Rich minimal (medium)
RNA-Seq	RNA sequencing approach
Str	Streptomycin
T3SS	Type III secretion system
Tn-Seq	Transposon mutagenesis sequencing
WT	Wild type

Chapter One

Literature Review

INTRODUCTION

There are multiple ways that bacteria can interact with their hosts, such as commensalism, mutualism and parasitism. Better understanding these interactions will provide the foundation to manipulate them for the benefit of humans, as well as the maintenance of a sustainable ecosystem. Quorum sensing (QS) is a mechanism that bacteria use to communicate with each other. It plays an important role in the host-bacteria interactions in a number of systems. *Pantoea stewartii* subspecies *stewartii* (*P. stewartii*), a phytopathogen causing wilt in corn, utilizes QS to control expression of genes that play a critical role during *in planta* virulence. Its exopolysaccharide, stewartan, is known as a key virulence factor controlled by the QS regulon. However, the function and integration of QS downstream targets in *P. stewartii* are poorly understood. In addition, the genes important for *in vitro* and *in planta* growth of this wilt-causing pathogen are understudied.

Pantoea stewartii subspecies *stewartii,* a corn pathogen

P. stewartii, a Gram-negative rod-shaped, gamma-proteobacterium belonging to the *Enterobacteriaceae* family containing plant-associate enterics (e.g. *Erwinia*, *Dickeya*, and *Pectobacterium* sp.) and plant-associated human enteric pathogens (e.g. *Escherichia* and *Salmonella* sp.), causes a bacterial infection, named Stewart's wilt, in maize (1). The disease causes a reduction in plant productivity and therefore economic loss. Sweet corn and popcorn varieties are the main host since resistant hybrid field corn varieties are available (2). *P. stewartii* is transmitted passively to the plant during the feeding process of corn flea beetles, *Chaetocnema pulicaria* (3). This second host helps to maintain the life cycle of this phytopathogen (Figure 1.1). The bacterium resides over winter in the gut of the flea beetles, and gains access into the plants through excrement deposited on the wounds generated by the insect scratching the plant tissues. The annual disease incidence is proportional to insect prevalence in the corn fields (1, 4)

as milder winters associated with global warming contribute to higher levels of disease in corn. In the first phase of the disease, visible water-soaked lesions appear on the infected leaves due to the effect of a type III secretion system (T3SS) effector, WtsE, which is delivered through the Hrp T3SS gene system (5). In the second phase of infection, *P. stewartii* migrates from the leaf apoplast into the xylem where it grows to form a biofilm that blocks water transport causing the wilt disease symptoms (1). Death can occur if the plants were infected at the seedling phase (6).

Other xylem-dwelling phytopathogens

In addition to *P. stewartii*, there are other xylem-colonizing pathogenic bacteria, such as *Erwinia amylovora*, *Pseudomonas syringae* pv. *actinidiae*, *Ralstonia solanacearum*, *Xanthomonas oryzae* pv. *oryzae*, and *Xylella fastidiosa*, that cause damage to commercially important crops. *E. amylovora* is a Gram-negative bacterium causing fire blight disease in *Rosaceae* plants (7). *P. syringae* pv. *actinidiae*, a Gram-negative bacterium, is a causal agent of bacterial canker in kiwi (8). *R. solanacearum*, a Gram-negative bacterium causing bacterial wilt disease in multiple plant species such as tomato, potato and pepper, can survive in soil for an extensive period (9). The Gram-negative bacterium *X. oryzae* pv. *oryzae* is the causative agent of the most serious disease of rice, bacterial blight (10). *X. fastidiosa* is a fastidious Gram-negative bacterium causing different diseases in woody plants, including citrus variegated chlorosis disease in citrus and Pierce's disease in grapevine (11). These pathogens usually gain access to the xylem tissues and cause disease symptoms by interfering with the normal functions of the xylem (12).

The xylem structure contains three main cell types: xylem tracheary elements, xylary fibers and xylem parenchyma cells (13). Both tracheary elements and xylary fibers are metabolically dead cells with extensive secondary cell wall formation to create and support the

xylem vessels. The parenchyma cells, located outside of the xylem vessels, are metabolically active (13). The main function of xylem vessels is water transport; thus, xylem sap contains a low concentration of nutrients, such as some organic acids, amino acids and sugars, to serve as a nutrient source for these xylem-dwelling bacterial pathogens (14). The xylem colonization can lead to biofilm formation and water blockage (15), which results in wilt and even death of the infected plants. QS regulation is known to coordinate the formation of biofilm at high cell density in a number of bacteria, including *P. stewartii* (16, 17).

Bacterial quorum sensing

QS is a cell-to-cell communication system found primarily in eubacteria. During QS, bacteria produce autoinducers as the QS signal, which are normally N-acyl homoserine lactones (AHL) or post-translationally processed peptides in Gram-negative proteobacteria and Gram-positive bacteria, respectively. In the AHL-dependent systems, the QS signal interacts with the master regulatory protein, a LuxR-family protein, in the bacterium to coordinate its gene expression in a cell-density dependent manner. At low cell density (LCD) the probability of interaction between the QS signal and its cognate regulatory protein is low compared to high cell density (HCD). Therefore, the gene expression at different cell densities can be differentially coordinated to the benefit of the bacteria. QS is known to control a number of various phenotypes, including bioluminescence, biofilm formation, extracellular enzyme secretion, and virulence (17, 18).

The bioluminescence system in *Vibrio fischeri*, a symbiont of the Hawaiian bobtail squid (*Euprymna scolopes*), is considered a paradigm of studying QS in Gram-negative proteobacteria (19). In *V. fischeri*, the LuxI autoinducer synthase produces the major diffusible signaling molecule 3-N-oxohexanoyl-L-homoserine lactone (3-oxo-C_6-HSL) (20). When the cell

population increases in density (in the light organ of the squid or in a batch culture in the laboratory), the binding of AHL to the transcriptional activator protein LuxR leads to the activation of the *luxICDABEG* genes which generate luminescence (21, 22). Other genes in the LuxR regulon are also activated (23). A second QS system has been identified in *V. fischeri* with AinS, the homolog of LuxI which is an N-octanoyl homoserine lactone (C8-HSL) synthase, and LitR, a transcriptional repressor that regulate gene expression of colonization factors (24). The two QS systems work cooperatively to establish colonization and luminescence in the host (24).

Homologues of LuxR are members of the LuxR protein subfamily whose functions are to detect and respond to AHL signals through the N-terminal domain with DNA-binding activity in the C-terminal domain controlling transcription of target genes. Based on the multimeric characteristics of LuxR proteins and the interactions between them and AHL, these homologues have been classified into five categories (19). Classes 1-3 comprise activators that only function upon interacting with the AHL molecules. *Agrobacterium tumefaciens* TraR is a typical member of class I regulators which bind AHL cotranslationally. AHL is included within the structure of TraR, which forms a dimer and functions at HCD, while TraR is a target of degradation by the Clp and Lon proteases at LCD (25). LuxR in *V. fischeri* is a member of category II activators that need AHL to stabilize the protein folding during translation and form inactive dimers at LCD but become more stable at higher AHL concentrations, enabling them to function at HCD in a dimeric form (26). Class III homologues, with *Mesorhizobium tianshanense* MrtR as an example, are activators that form stable monomeric unfunctional protein at LCD but interact with AHL at HCD to form active dimers (27). On the contrary, group IV regulators comprise just a few members capable of functioning as repressors or activators in their dimeric form, which occurs in the absence of AHL at LCD. *P. stewartii* EsaR belongs to this category of LuxR family

proteins (28, 29). The last category of LuxR homologues consist of "orphan" LuxR proteins with the absence of a cognate LuxI homologue in the same cell. Proteins in this group do not dimerize, even when interacting with AHL, like SdiA from *Escherichia coli* (30).

Quorum sensing and *P. stewartii*

P. stewartii uses the same QS signal, 3-oxo-C6-HSL, as *Vibrio fischeri*. This signal is synthesized by the AHL synthase EsaI, a LuxI homologue (31). In *P. stewartii*, the master QS regulatory protein, EsaR, a LuxR homologue (31), is active and binds to DNA regions to repress (32) or activate (33) different gene targets at LCD. However, when EsaR forms complexes with AHL at HCD, it becomes inactive and unable to bind to DNA to control its targets (Figure 1.2) (34).

EsaR is known to autorepress its own gene (35). It has been also established that the major virulence factor of *P. stewartii*, the stewartan exopolysaccharide, is controlled by quorum sensing (32). EsaR represses the expression of *rcsA* encoding a transcriptional activator of capsule production (28). External addition of AHL in the medium resulted in a mucoid phenotype of the AHL synthase EsaI mutant, which exhibited a "dry" appearance in the absence of AHL. Whereas the *rcsA* mutant displayed a "dry" phenotype in the absence or presence of external AHL. This suggested that the *P. stewartii* QS controls capsule production via RcsA activity. Minogue et al. also showed that this control is mediated via the direct binding of EsaR with the *rcsA* promoter rather than any of the downstream targets of RcsA such as *wceG*, *wza*, *wceL* and *wceB* that are required for capsule synthesis (28). RcsA is an unstable protein in the Rcs (regulation of capsule synthesis) regulatory network involving in colanic acid and K antigen synthesis in *Escherichia coli* (36, 37) and *Salmonella typhi* (38) as well as capsule production in *P. stewartii* and *Erwinia amylovora* (39). Moreover, proteomic and transcriptomic analysis have

revealed that EsaR directly controls three distinct physiological functions of *P. stewartii*: stress response, cell wall and capsule synthesis, and surface motility and adhesion as well as other genes, including *dkgA,* encoding a 2,5-diketo gluconate reductase, and *esaS,* encoding a small RNA (40, 41).

Capsule production in *P. stewartii*

At HCD, when EsaR is inactive and RcsA is highly expressed, *P. stewartii* produces a large amount of stewartan exopolysaccharide (EPS) as a key virulence factor causing the wilt disease in plants (42). Stewartan EPS is an anionic polymer composed of heptasaccharide (with D-glucose, D-galactose, and D-glucuronic acid in a 3:3:1 molar ratio) repeat units (43). This high-molecular-weight heteropolysaccharide is classified as a group 1 polysaccharide and regulated by a RcsC/YojN/RcsB/A phosphorelay signal transduction system (39, 44). The expression of stewartan in *P. stewartii* is controlled by QS via the repression of *rcsA* as discussed above. Specifically, RcsA expression is directly repressed by EsaR in the absence of AHL (28, 32). As RcsA levels rise at HCD, it controls EPS expression by activating the capsule biosynthesis genes (*cps* or now *wce*) (28, 45). In *P. stewartii*, stewartan production is dependent on three clusters of *wce* genes, *wce-I, wce-II* and *wce-III,* and RcsA controls the expression of all of these gene clusters (46). Mutants from the *wce-II* cluster (*wceO* and *wzx2*) are incapable of stewartan EPS biosynthesis while a *wce-III* (contain a single gene *wceG2* with redundant function to *wceG1*) mutant produces less stewartan compared to the wild-type strain (46). Studies in the Stevens laboratory have demonstrated that EsaR also directly controls some of the *wce* genes, such as *wceL* and *wceG2* (40), which is contrary to the notion that QS exclusively controls stewartan production only indirectly via RcsA regulation (28, 46). This indicates the presence of a coherent feed forward loop that allows rapid response to the QS signal (47).

Surface motility in *P. stewartii*

Bacterial movements can be performed by a variety of different mechanisms such as swarming, swimming, twitching, gliding and sliding motility (48, 49). Both swarming and swimming-based motility are flagella-dependent movement. Swimming is an individual activity that takes place in liquid, while swarming is a group movement across a surface. Twitching motility is another type of surface movement powered by type IV pili extension and retraction. Gliding motility is most commonly independent of either flagella or pili. The last type of motility is sliding, which is a passive form of surface movement depending on surfactants allowing the colony to spread away from the center. Motility is important for biofilm formation (50, 51) and it provides the bacteria some advantages in nutrient scavenging as well as invasion (48, 49). Therefore, it contributes at least partially to the virulence and/or survival of bacteria in a specific niche.

Swarming motility is currently found to be restricted in three families of bacteria: *Gammaproteobacteria*, *Alphaproteobacteria* and *Firmicutes* (49). This restriction could be biased by the use of laboratory conditions which may not facilitate the swarming behavior (52). Swarming usually requires an energy-rich medium to provide energy for flagellar activity, solid medium with above 0.3% agar concentration (below 0.3% is considered swimming), cell-to-cell interaction, and surfactant synthesis and secretion (49).

It has been demonstrated that *P. stewartii* possesses swarming motility which is critical to the pathogenicity of the bacterium (53). This motility is promoted by glucose supplemented in the growth medium to provide energy. Sucrose (a phosphoenolpyruvate phosphotransferase system, PTS, sugar) and galactose (a non-PTS sugar) partially support the motility whereas mannitol (a PTS sugar), and lactose and glycerol (non-PTS sugars) do not support swarming

(53). Addition of potassium, sodium and carbonyl cyanide 3-chlorophenylhydrazone inhibits the motility as well. In *P. stewartii*, QS controls this motility in an interesting and complicated manner: a lack of AHL inhibits the motility, while lacking both AHL and EsaR attenuates swarming (53). This suggests that EsaR inhibits surface motility at LCD whereas surface motility is promoted at HCD. It has been proposed that QS indirectly controls surface motility via stewartan synthesis controlled by the RcsA/B phosphorelay regulatory system (53). In addition, it is known that expression of the $FlhD_2C_2$ flagellar master regulator is directly control by the RcsA/B system in other bacteria (54-56). Finally, it has been shown in *P. stewartii* that siderophore-mediated iron acquisition is essential for surface motility *in vitro* and bacterial movement *in planta* (57).

Proteomic and transcriptomic data from our lab showed that expression of the LysR-type regulator LrhA is controlled directly by EsaR, the master regulator of QS in *P. stewartii* (40, 41). LrhA was shown to be the key regulator controlling the expression of flagella, motility and chemotaxis in *E. coli* by regulating the synthesis and concentration of $FlhD_2C_2$, the master regulator of flagella and chemotaxis gene expression (58). In *E. coli*, LrhA positively autoregulates its own expression and represses the expression of *flhDC*; thereby, suppressing motility and chemotaxis (58). The role of LrhA in *P. stewartii* was examined as part of this dissertation using a RNA-Seq and follow-up molecular-based approaches (Chapters Two and Four).

Next generation sequencing and bioinformatic tools to study bacterial transcriptomes and host-microbe interactions

The development of next generation sequencing (NGS) and bioinformatics has advanced the approaches used in biological research. The massive amount of information generated by

NGS can be analyzed using bioinformatic tools to understand biological processes at the genome as well as transcriptome levels (59-61). Illumina high-throughput sequencing is one of the most popular NGS that provides DNA sequencing information by synthesis in forms of short-length reads in the range of 50-300 bp (http://www.illumina.com). In Illumina sequencing, DNA samples are processed to have a certain length including specific adapters at the ends of the DNA fragments to facilitate their attachment to oligonucleotides on the Illumina chip for clonal amplification of the original DNA sequences. Then fluorescently labeled reversible ddNTPs are added to the chip during the sequencing process. An image of the entire chip containing millions of clonal clusters of different DNA templates in parallel is taken every time a round of the labelled nucleotides are added to the primers that anneal to the templates. Each nucleotide has its own fluorescent signal to distinguish among them. After that, these ddNTPs are modified to allow incorporation of the next ddNTPs to repeat the sequencing cycles. In this way, the sequences of the original templates are deciphered by the process known as sequencing by synthesis. An overall workflow of Illumina sequencing is demonstrated in Figure 1.3. In this dissertation, Illumina sequencing was utilized in multiple approaches such as RNA-Seq (Chapter Two), whole genome sequencing (Chapter Three) and Tn-Seq (Chapter Five) to reveal the downstream targets of the *P. stewartii* QS regulon and to identify genes important for *in planta* colonization and survival of this phytopathogen.

During RNA-Seq, the bacterial transcriptome is extracted from tested conditions and converted to complementary DNA (cDNA) for high-throughput sequencing. The raw data is then mapped to a reference to evaluate the number of transcripts throughout the genome and to compare gene expression under different conditions. This analysis provides information about up- or down-regulated genes at different growth conditions for one type of organism as well as

differential regulation by a certain gene when comparing between a mutant and its parental wild-type strain under the same growth condition.

In whole genome sequencing, Illumina sequencing provides high-coverage of sequencing data which leads to the high accuracy of genome assembly. One of the major advances of Illumina NGS technology is the usage of paired-end (PE) sequencing. Both ends of the DNA fragments in a sequencing library are sequenced and aligned as read pairs during PE sequencing. Since the distance between two ends of the DNA fragments is pre-determined during the fragmentation process, PE sequencing allows for the generation of a higher number of reads for each library preparation compared to single-read data. In addition, PE sequencing provides better alignment which can overcome the problem of repetitive sequences in the templates (62). Coupling with PE sequencing technology is the Illumina mate-pair (MP) library preparation process to facilitate the whole genome sequencing of a complex prokaryotic genome such as *P. stewartii*, containing multiple plasmids and a large number of repetitive sequences like transposons and insertion sequences. During MP library preparation, the ends of DNA fragments are repaired with labeled dNTPs after DNA fragmentation to a relatively large size (usually greater than 1 kilobase fragments). Then, the DNA molecules are circularized prior to a second round of fragmentation to a smaller size that is suitable for sequencing on an Illumina sequencer platform. The labeled fragments, corresponding to the ends of the original DNA fragment, are affinity-purified and processed to be sequenced with PE sequencing technology. This allows for the longer distance between two mates of a pair from the original DNA templates to be accurately sequenced and then aligned to overcome the problem caused by the large repetitive sequences. This technology has been successfully used to complete the assembly of the organism of interest, *P. stewartii*, further discussed in Chapter Three of this dissertation.

Finally, Tn-Seq is also a product of NGS in which a library of transposon mutants is generated and tested under different conditions prior to collecting DNA for high-throughput sequencing. The positions of transposons are located and the numbers of transposons within each gene are counted and compared with the same gene at a different condition. For example, if the number of the insertions in gene A is lower in the treatment compared to in the control, it means that gene A may play an important role in the survival of the organism in the treatment in comparison to the control. This approach is similar to reverse genetics in which the comparison of the phenotype of the wild-type and mutant strain is made to decipher the role of the mutant gene; however, Tn-Seq provides multiple comparisons of the entire genome simultaneously.

RESEARCH PLAN

It has previously been shown that QS plays a critical role in the pathogenesis of *P. stewartii* in corn (1, 16). However, the regulon downstream of EsaR is less well understood. Chapters Two and Four describe efforts to elucidate the function and integration of two direct targets of EsaR: LrhA and RcsA. Both of these proteins are transcription factors whose regulons were investigated using RNA-Seq (Chapter Two). A regulatory connection between them was identified in Chapter Two and confirmed in Chapter Four. Chapter Three describes the complete assembly of the *P. stewartii* genome which provided insights into a missing 66-kilobase region of chromosomal DNA, further discussed in Chapter Four. The complete genome assembly revealed the location of the two T3SS important for host-association in different niches (*in planta* vs. in corn flea beetle) on two mega-plasmids. It also has facilitated the analysis of a global mutagenesis using a mariner-based transposon to study genes important for *in vitro* and *in planta* growth via a Tn-Seq approach described in Chapter Five. These studies have extended our

knowledge about this phytopathogen and may eventually lead to an informed strategy to control this bacterial wilt disease.

REFERENCES

1. **Roper MC.** 2011. *Pantoea stewartii* subsp. *stewartii*: lessons learned from a xylem-dwelling pathogen of sweet corn. Molecular Plant Pathology **12:**628-637.
2. **Pataky J.** 2003. Stewart's Wilt of Corn. APSnet Features. Online.
3. **Esker PD, Nutter FW.** 2002. Assessing the risk of Stewart's disease of corn through improved knowledge of the role of the corn flea beetle vector. Phytopathology **92:**668-670.
4. **Menelas B, Block CC, Esker PD, Nutter FW, Jr.** 2006. Quantifying the feeding periods required by corn flea beetles to acquire and transmit *Pantoea stewartii*. Plant Disease **90:**319-324.
5. **Ham JH, Majerczak DR, Arroyo-Rodriguez AS, Mackey DM, Coplin DL.** 2006. WtsE, an AvrE-family effector protein from *Pantoea stewartii* subsp. *stewartii*, causes disease-associated cell death in corn and requires a chaperone protein for stability. Molecular Plant-Microbe Interactions **19:**1092-1102.
6. **Braun EJ.** 1982. Ultrastructural investigation of resistant and susceptible maize inbreds infected with *Erwinia stewartii*. Phytopathology **72:**159-166.
7. **Eastgate JA.** 2000. *Erwinia amylovora:* the molecular basis of fireblight disease. Molecular Plant Pathology **1:**325-329.
8. **Scortichini M, Marcelletti S, Ferrante P, Petriccione M, Firrao G.** 2012. *Pseudomonas syringae* pv. *actinidiae*: a re-emerging, multi-faceted, pandemic pathogen. Molecular Plant Pathology **13:**631-640.
9. **Genin S, Boucher C.** 2002. *Ralstonia solanacearum*: secrets of a major pathogen unveiled by analysis of its genome. Molecular Plant Pathology **3:**111-118.
10. **Nino-Liu DO, Ronald PC, Bogdanove AJ.** 2006. *Xanthomonas oryzae* pathovars: model pathogens of a model crop. Molecular Plant Pathology **7:**303-324.
11. **Mang SM, Frisullo S, Elshafie HS, Camele I.** 2016. Diversity evaluation of *Xylella fastidiosa* from infected olive trees in Apulia (Southern Italy). The Plant Pathology Journal **32:**102-111.
12. **Mansfield J, Genin S, Magori S, Citovsky V, Sriariyanum M, Ronald P, Dow M, Verdier V, Beer SV, Machado MA, Toth I, Salmond G, Foster GD.** 2012. Top 10 plant pathogenic bacteria in molecular plant pathology. Molecular Plant Pathology **13:**614-629.
13. **Schuetz M, Smith R, Ellis B.** 2013. Xylem tissue specification, patterning, and differentiation mechanisms. Journal of Experimental Botany **64:**11-31.
14. **Fatima U, Senthil-Kumar M.** 2015. Plant and pathogen nutrient acquisition strategies. Frontiers in Plant Science **6:**750.
15. **Bae C, Han SW, Song YR, Kim BY, Lee HJ, Lee JM, Yeam I, Heu S, Oh CS.** 2015. Infection processes of xylem-colonizing pathogenic bacteria: possible explanations for the scarcity of qualitative disease resistance genes against them in crops. Theoretical and Applied Genetics. **128:**1219-1229.
16. **von Bodman SB, Bauer WD, Coplin DL.** 2003. Quorum sensing in plant-pathogenic bacteria. Annual Review of Phytopathology **41:**455-482.
17. **Rutherford ST, Bassler BL.** 2012. Bacterial quorum sensing: its role in virulence and possibilities for its control. Cold Spring Harbor Perspectives in Medicine **2**:a012427
18. **de Kievit TR, Iglewski BH.** 2000. Bacterial quorum sensing in pathogenic relationships. Infection and Immunity **68:**4839-4849.
19. **Stevens AM, Queneau Y, Soulere L, von Bodman S, Doutheau A.** 2011. Mechanisms and synthetic modulators of AHL-dependent gene regulation. Chemical Reviews **111:**4-27.
20. **Eberhard A, Burlingame AL, Eberhard C, Kenyon GL, Nealson KH, Oppenheimer NJ.** 1981. Structural identification of autoinducer of *Photobacterium fischeri* luciferase. Biochemistry **20:**2444-2449.

21. **Engebrecht J, Silverman M.** 1984. Identification of genes and gene products necessary for bacterial bioluminescence. Proceedings of the National Academy of Sciences of the United States of America **81:**4154-4158.
22. **Meighen EA.** 1991. Molecular biology of bacterial bioluminescence. Microbiological Reviews **55:**123-142.
23. **Antunes LC, Schaefer AL, Ferreira RB, Qin N, Stevens AM, Ruby EG, Greenberg EP.** 2007. Transcriptome analysis of the *Vibrio fischeri* LuxR-LuxI regulon. Journal of Bacteriology **189:**8387-8391.
24. **Lupp C, Ruby EG.** 2005. *Vibrio fischeri* uses two quorum-sensing systems for the regulation of early and late colonization factors. Journal of Bacteriology **187:**3620-3629.
25. **Zhu J, Winans SC.** 2001. The quorum-sensing transcriptional regulator TraR requires its cognate signaling ligand for protein folding, protease resistance, and dimerization. Proceedings of the National Academy of Sciences of the United States of America **98:**1507-1512.
26. **Urbanowski ML, Lostroh CP, Greenberg EP.** 2004. Reversible acyl-homoserine lactone binding to purified *Vibrio fischeri* LuxR protein. Journal of Bacteriology **186:**631-637.
27. **Yang M, Giel JL, Cai T, Zhong Z, Zhu J.** 2009. The LuxR family quorum-sensing activator MrtR requires its cognate autoinducer for dimerization and activation but not for protein folding. Journal of Bacteriology **191:**434-438.
28. **Minogue TD, Carlier AL, Koutsoudis MD, von Bodman SB.** 2005. The cell density-dependent expression of stewartan exopolysaccharide in *Pantoea stewartii* ssp. *stewartii* is a function of EsaR-mediated repression of the *rcsA* gene. Molecular Microbiology **56:**189-203.
29. **Schu DJ, Carlier AL, Jamison KP, von Bodman S, Stevens AM.** 2009. Structure/function analysis of the *Pantoea stewartii* quorum-sensing regulator EsaR as an activator of transcription. Journal of Bacteriology **191:**7402-7409.
30. **Patankar AV, Gonzalez JE.** 2009. Orphan LuxR regulators of quorum sensing. FEMS Microbiology Reviews **33:**739-756.
31. **Beck von Bodman S, Farrand SK.** 1995. Capsular polysaccharide biosynthesis and pathogenicity in *Erwinia stewartii* require induction by an N-acylhomoserine lactone autoinducer. Journal of Bacteriology **177:**5000-5008.
32. **von Bodman SB, Majerczak DR, Coplin DL.** 1998. A negative regulator mediates quorum-sensing control of exopolysaccharide production in *Pantoea stewartii* subsp. *stewartii*. Proceedings of the National Academy of Sciences of the United States of America **95:**7687-7692.
33. **von Bodman SB, Ball JK, Faini MA, Herrera CM, Minogue TD, Urbanowski ML, Stevens AM.** 2003. The quorum sensing negative regulators EsaR and ExpR(Ecc), homologues within the LuxR family, retain the ability to function as activators of transcription. Journal of Bacteriology **185:**7001-7007.
34. **Shong J, Huang YM, Bystroff C, Collins CH.** 2013. Directed evolution of the quorum-sensing regulator EsaR for increased signal sensitivity. ACS Chemical Biology **8:**789-795.
35. **Minogue TD, Wehland-von Trebra M, Bernhard F, von Bodman SB.** 2002. The autoregulatory role of EsaR, a quorum-sensing regulator in *Pantoea stewartii* ssp. *stewartii*: evidence for a repressor function. Molecular Microbiology **44:**1625-1635.
36. **Ebel W, Trempy JE.** 1999. *Escherichia coli* RcsA, a positive activator of colanic acid capsular polysaccharide synthesis, functions to activate its own expression. Journal of Bacteriology **181:**577-584.
37. **Keenleyside WJ, Jayaratne P, MacLachlan PR, Whitfield C.** 1992. The *rcsA* gene of *Escherichia coli* O9:K30:H12 is involved in the expression of the serotype-specific group I K (capsular) antigen. Journal of Bacteriology **174:**8-16.

38. **Virlogeux I, Waxin H, Ecobichon C, Lee JO, Popoff MY.** 1996. Characterization of the *rcsA* and *rcsB* genes from *Salmonella typhi: rcsB* through *tviA* is involved in regulation of Vi antigen synthesis. Journal of Bacteriology **178:**1691-1698.
39. **Wehland M, Kiecker C, Coplin DL, Kelm O, Saenger W, Bernhard F.** 1999. Identification of an RcsA/RcsB recognition motif in the promoters of exopolysaccharide biosynthetic operons from *Erwinia amylovora* and *Pantoea stewartii* subspecies *stewartii*. The Journal of Biological Chemistry **274:**3300-3307.
40. **Ramachandran R, Burke AK, Cormier G, Jensen RV, Stevens AM.** 2014. Transcriptome-based analysis of the *Pantoea stewartii* quorum-sensing regulon and identification of EsaR direct targets. Applied and Environmental Microbiology **80:**5790-5800.
41. **Ramachandran R, Stevens AM.** 2013. Proteomic analysis of the quorum-sensing regulon in *Pantoea stewartii* and identification of direct targets of EsaR. Applied and Environmental Microbiology **79:**6244-6252.
42. **Bradshaw-Rouse JJ, Whatley MH, Coplin DL, Woods A, Sequeira L, Kelman A.** 1981. Agglutination of *Erwinia stewartii* strains with a corn agglutinin: correlation with extracellular polysaccharide production and pathogenicity. Applied and Environmental Microbiology **42:**344-350.
43. **Nimtz M, Mort A, Wray V, Domke T, Zhang Y, Coplin DL, Geider K.** 1996. Structure of stewartan, the capsular exopolysaccharide from the corn pathogen *Erwinia stewartii*. Carbohydrate Research **288:**189-201.
44. **Takeda S, Fujisawa Y, Matsubara M, Aiba H, Mizuno T.** 2001. A novel feature of the multistep phosphorelay in *Escherichia coli*: a revised model of the RcsC --> YojN --> RcsB signalling pathway implicated in capsular synthesis and swarming behaviour. Molecular Microbiology **40:**440-450.
45. **Reeves PR, Hobbs M, Valvano MA, Skurnik M, Whitfield C, Coplin D, Kido N, Klena J, Maskell D, Raetz CR, Rick PD.** 1996. Bacterial polysaccharide synthesis and gene nomenclature. Trends in Microbiology **4:**495-503.
46. **Carlier A, Burbank L, von Bodman SB.** 2009. Identification and characterization of three novel EsaI/EsaR quorum-sensing controlled stewartan exopolysaccharide biosynthetic genes in *Pantoea stewartii* ssp. *stewartii*. Molecular Microbiology **74:**903-913.
47. **Mangan S, Zaslaver A, Alon U.** 2003. The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks. Journal of Molecular Biology **334:**197-204.
48. **Harshey RM.** 2003. Bacterial motility on a surface: many ways to a common goal. Annual Review of Microbiology **57:**249-273.
49. **Kearns DB.** 2010. A field guide to bacterial swarming motility. Nature Reviews. Microbiology **8:**634-644.
50. **Houry A, Briandet R, Aymerich S, Gohar M.** 2010. Involvement of motility and flagella in *Bacillus cereus* biofilm formation. Microbiology **156:**1009-1018.
51. **Ramsey MM, Whiteley M.** 2004. *Pseudomonas aeruginosa* attachment and biofilm development in dynamic environments. Molecular Microbiology **53:**1075-1087.
52. **Patrick JE, Kearns DB.** 2009. Laboratory strains of *Bacillus subtilis* do not exhibit swarming motility. Journal of Bacteriology **191:**7129-7133.
53. **Herrera CM, Koutsoudis MD, Wang X, von Bodman SB.** 2008. *Pantoea stewartii* subsp. *stewartii* exhibits surface motility, which is a critical aspect of Stewart's wilt disease development on maize. Molecular Plant-Microbe Interactions **21:**1359-1370.
54. **Francez-Charlot A, Laugel B, Van Gemert A, Dubarry N, Wiorowski F, Castanie-Cornet MP, Gutierrez C, Cam K.** 2003. RcsCDB His-Asp phosphorelay system negatively regulates the *flhDC* operon in *Escherichia coli*. Molecular Microbiology **49:**823-832.

55. **Clemmer KM, Rather PN.** 2007. Regulation of *flhDC* expression in *Proteus mirabilis*. Research in Microbiology **158:**295-302.
56. **Wang Q, Zhao Y, McClelland M, Harshey RM.** 2007. The RcsCDB signaling system and swarming motility in *Salmonella enterica* serovar *typhimurium*: dual regulation of flagellar and SPI-2 virulence genes. Journal of Bacteriology **189:**8447-8457.
57. **Burbank L, Mohammadi M, Roper MC.** 2014. Siderophore-mediated iron acquisition influences motility and is required for full virulence for the xylem-dwelling bacterial phytopathogen, *Pantoea stewartii* subsp. *stewartii*. Applied and Environmental Microbiology.
58. **Lehnen D, Blumer C, Polen T, Wackwitz B, Wendisch VF, Unden G.** 2002. LrhA as a new transcriptional key regulator of flagella, motility and chemotaxis genes in *Escherichia coli*. Molecular Microbiology **45:**521-532.
59. **Levy SE, Myers RM.** 2016. Advancements in next-generation sequencing. Annual Review of Genomics and Human Genetics **17:**95-115.
60. **Pareek CS, Smoczynski R, Tretyn A.** 2011. Sequencing technologies and genome sequencing. Journal of Applied Genetics **52:**413-435.
61. **Martin JA, Wang Z.** 2011. Next-generation transcriptome assembly. Nature Reviews. Genetics **12:**671-682.
62. **Nakazato T, Ohta T, Bono H.** 2013. Experimental design-based functional mining and characterization of high-throughput sequencing data in the sequence read archive. PLOS ONE **8:**e77910.

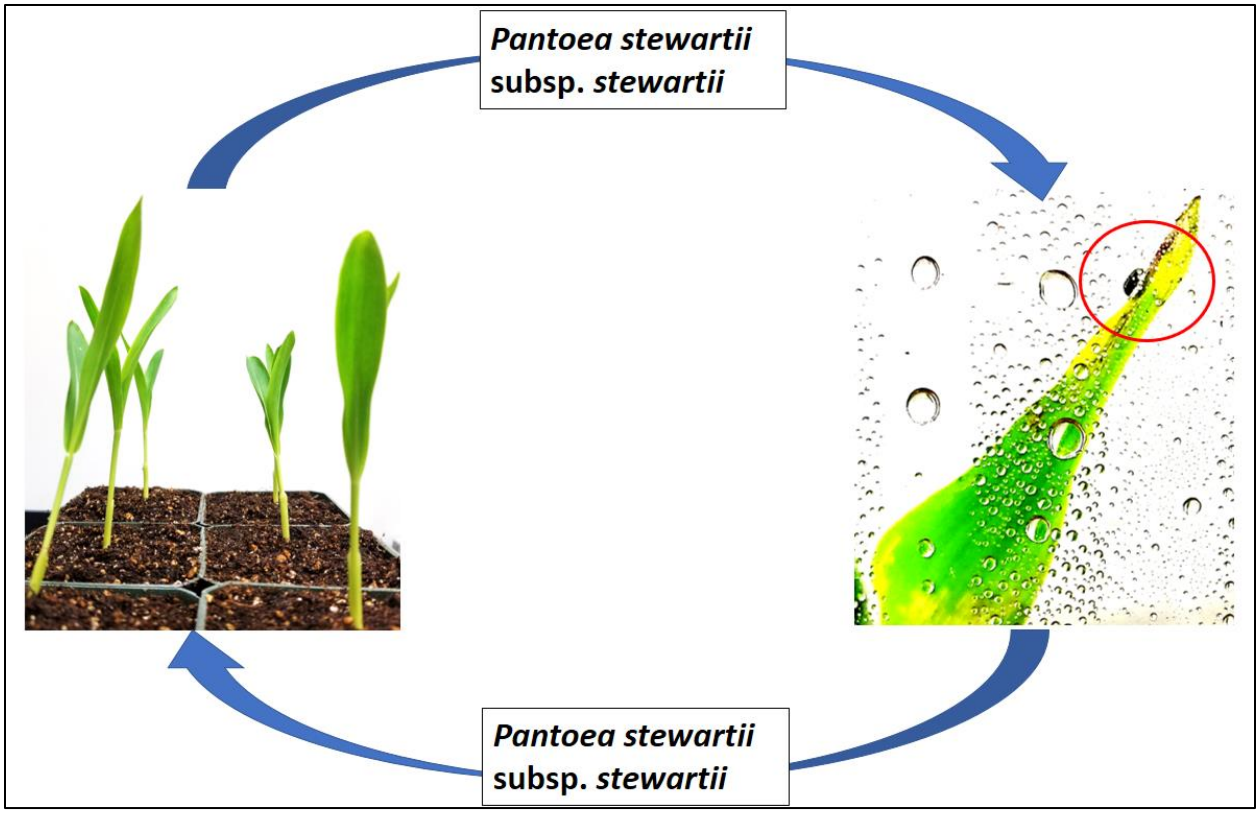

Figure 1.1. Schematic life cycle of the phytopathogen *P. stewartii* subsp. *stewartii*. The *P. stewartii* bacterium resides in the gut of the corn flea beetle (highlighted in the red circle in the left) and which transmits it to the corn during the insect's feeding process. If naïve beetles feed on infected plants, they become colonized and can then spread the bacterium to other corn plants. The beetles help to maintain the bacterial survival over the winter inside their guts.

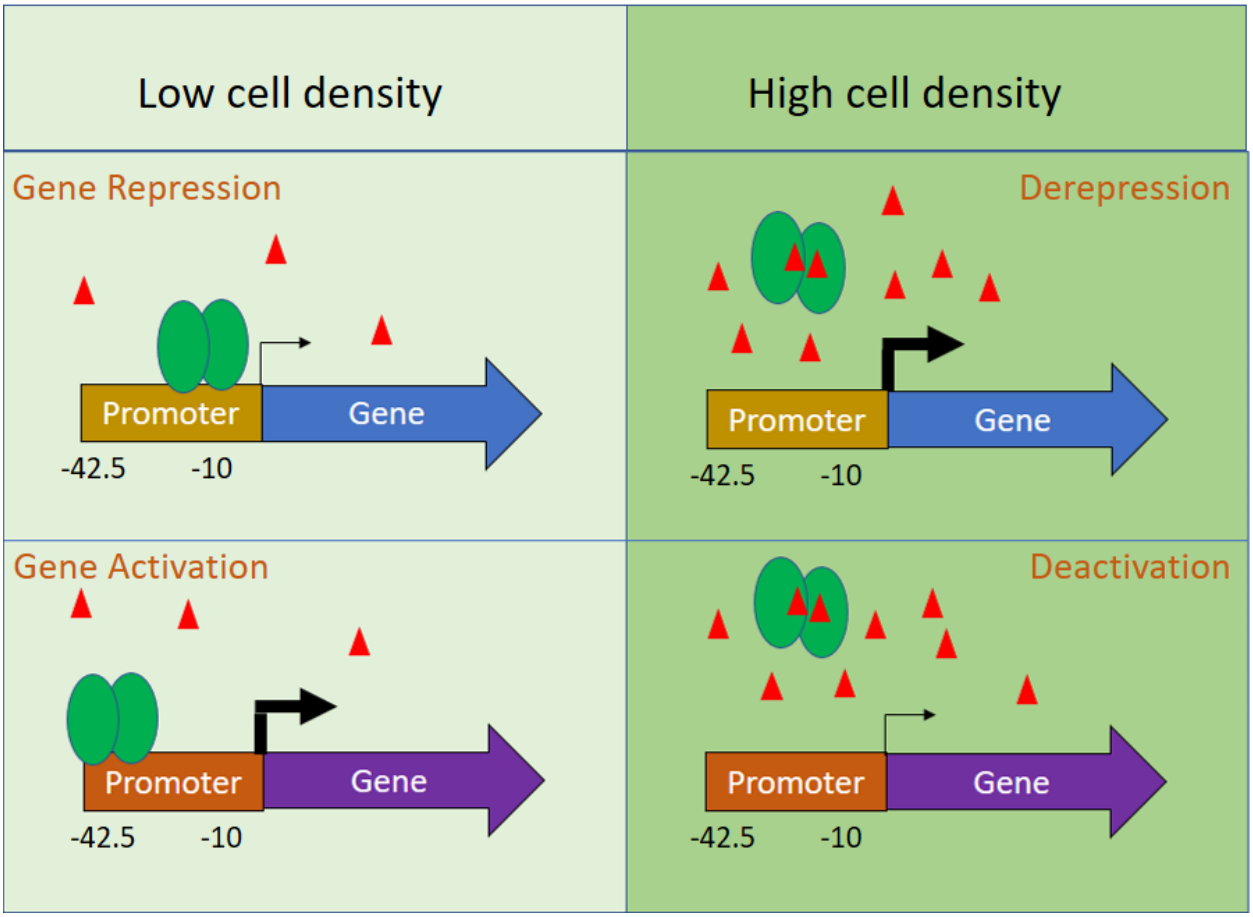

Figure 1.2. Schematic diagram of quorum sensing in the phytopathogen *P. stewartii* subsp. *stewartii*. At low cell density, EsaR (green ovals) forms a homodimer and interacts with promoter regions of its targets to repress (top) or activate (bottom) the transcription of these genes. Acyl homoserine lactones (AHL, red triangle) forms complexes with EsaR at high cell density and prevents the EsaR from binding to the DNA, resulting in derepression or deactivation.

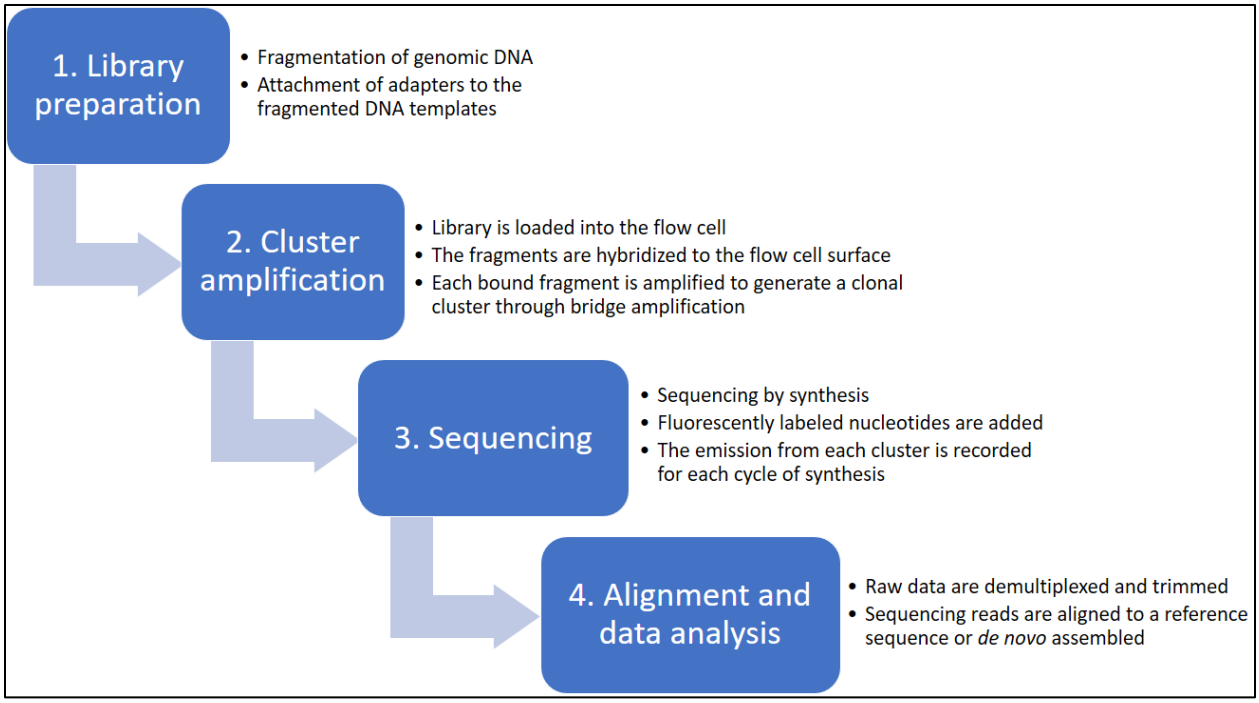

Figure 1.3. The general workflow of Illumina next-generation sequencing. This is the overall steps used in the generation of massive parallel high-throughput sequencing data using Illumina sequencing technology. The DNA templates used in each approach (RNA-Seq (Chapter Two), whole genome sequencing (Chapter Three), and Tn-Seq (Chapter Five)) are appropriately collected and specifically prepared during the library preparation step. Both RNA-Seq and Tn-Seq approaches in this dissertation used single end sequencing technology in step three of the workflow. The whole genome sequencing project utilized mate pair library preparation and pair end sequencing technology. All alignment and data analysis were done with Geneious bioinformatic software.

Chapter Two

Analyzing the Transcriptomes of Two Quorum-Sensing Controlled Transcription Factors, RcsA and LrhA, Important for *Pantoea stewartii* Virulence

Alison Kernell Burke[†]**, Duy An Duong**[†]**, Roderick V. Jensen,** and **Ann M. Stevens**[*]. 2015. Analyzing the transcriptomes of two quorum-sensing controlled transcription factors, RcsA and LrhA, important for *Pantoea stewartii* virulence. PLOS ONE 10(12): e0145358. https://doi.org/10.1371/journal.pone.0145358

[†] These authors contributed equally to the work.

*Corresponding author:
219 Life Sciences 1 (0910)
970 Washington St. SW
Virginia Tech
Blacksburg, VA 24061
Phone (540)-231-9378
FAX (540)-231-4043
Email ams@vt.edu

Key words: quorum sensing, RcsA, LrhA, *Pantoea stewartii*, transcriptome analysis, phytopathogen

Duy An Duong performed experiments illustrated in Table 2.3 and Figure 2.4.

Duy An Duong contributed with others to the data analysis and/or experiments for Tables 2.1, A.1, and A.2 and Figures 2.1, 2.2, 2.5, 2.6, 2.7.

ABSTRACT

The Gram-negative proteobacterium *Pantoea stewartii* subsp. *stewartii* causes wilt disease in corn plants. Wilting is primarily due to bacterial exopolysaccharide (EPS) production that blocks water transport in the xylem during the late stages of infection. EsaR, the master quorum-sensing (QS) regulator in *P. stewartii*, modulates EPS levels. At low cell densities EsaR represses or activates expression of a number of genes in the absence of its acyl homoserine lactone (AHL) ligand. At high cell densities, binding of AHL inactivates EsaR leading to derepression or deactivation of its direct targets. Two of these direct targets are the key transcription regulators RcsA and LrhA, which in turn control EPS production and surface motility/adhesion, respectively. In this study, RNA-Seq was used to further examine the physiological impact of deleting the genes encoding these two second-tier regulators. Quantitative reverse transcription PCR (qRT-PCR) was used to validate the regulation observed in the RNA-Seq data. A GFP transcriptional fusion reporter confirmed the existence of a regulatory feedback loop in the system between LrhA and RcsA. Plant virulence assays carried out with *rcsA* and *lrhA* deletion and complementation strains demonstrated that both transcription factors play roles during establishment of wilt disease in corn. These efforts further define the hierarchy of the QS-regulated network controlling plant virulence in *P. stewartii*.

INTRODUCTION

Pantoea stewartii subsp. *stewartii* (*P. stewartii*) is a Gram-negative rod-shaped, gamma-proteobacterium that belongs to the *Enterobacteriaceae* family containing plant-associate enterics (eg. *Erwinia*, *Dickeya*, and *Pectobacterium* spp.) and human enteric pathogens associated with plants (eg. *Escherichia* and *Salmonella* spp.). *P. stewartii* is the causative agent of Stewart's wilt in maize (1). It is transmitted passively from the gut of the corn flea beetle, *Chaetocnema pulicaria*, to the plant leaves via the feeding process of the insect vector (2). The symptoms of the disease include water-soaked lesions in the leaf when the bacterium is in the apoplast during the early stages of infection and wilt due to biofilm formation in the xylem during the later stages of infection; ultimately this can lead to plant death (3). The virulence of the bacterium is controlled through quorum sensing (QS) (4), a cell-to-cell communication system found primarily in eubacteria.

Gamma-proteobacteria, like *P. stewartii*, most commonly produce N-acyl homoserine lactone (AHL) signals during QS due to the activity of a LuxI-type protein. This intercellular signal then interacts with the master regulatory protein in the bacterium, a LuxR-type protein, to coordinate global gene expression. A variety of physiological outputs are controlled by QS such as bioluminescence, biofilm formation, virulence factor expression or exoenzyme production (5-9). In *P.* stewartii, the QS signal N-3-oxo-hexanoyl-L-homoserine lactone (3-oxo-C_6-HSL) is synthesized by the AHL synthase EsaI, a LuxI homologue (10). At low cell density, the master QS regulatory protein EsaR, a LuxR homologue, is active and binds its recognition sites in the DNA to repress or activate transcription of different gene targets (10-12). However, when EsaR and AHL form a complex, EsaR becomes inactive and unable to bind to DNA resulting in derepression or deactivation of target gene transcription at high cell density (13). Several direct

targets of EsaR have been identified through classic genetic (4), proteome-level (14) and transcriptome-level (15) analysis. Two of these direct targets, *rcsA* and *lrhA*, encode transcription factors elucidating additional levels of downstream regulation in response to QS.

RcsA is a transcription factor in the Rcs (regulation of capsule synthesis) regulatory network involved in colanic acid and K antigen capsular polysaccharide synthesis (cps) in *Escherichia coli* (16, 17) and *Salmonella typhi* (18). It has been established that a major virulence factor of *P. stewartii*, the exopolysaccharide (EPS), is controlled via direct EsaR-mediated repression of *rcsA* (CKS_2570) (4). Avirulent strains lacking EPS production typically have disruptions in genes found in either *rcsA* or the *cps* locus (4). In *P. stewartii*, RcsA (in conjunction with RcsB) has been proposed to activate three *cps/wce* gene clusters, comprised of (I) *wceG1, wza, wzb, wzc, wceL, wceB, wceM, wceN, wceF, wceJ, wceK, wzx, galF,* and *galE,* (II) *wceG2* and (III) *wceO and wzx* (4, 19). However, the complete RcsA regulon in *P. stewartii* has not been fully defined.

The LysR-type regulator LrhA (CKS_2075) is directly activated by EsaR in *P. stewartii* (14, 15). In *E. coli*, LrhA indirectly controls expression of flagella, motility and chemotaxis by positively autoregulating its own expression and repressing the synthesis of the master regulator of flagella and chemotaxis gene expression, the $FlhD_2C_2$ heterotetramer, thereby suppressing motility and chemotaxis (20). *P. stewartii* possesses swarming rather than swimming motility, and swarming motility is critical to the pathogenicity of the bacterium (21). QS controls this motility in an interesting and complicated manner; a lack of AHL inhibits the motility, while lacking both AHL and EsaR attenuates swarming. It has been proposed that QS indirectly controls motility via stewartan synthesis regulated by the RcsA/B phosphorelay system (21). Expression of the $FlhD_2C_2$ flagellar master regulator is directly control by the RcsA/B system in

other bacteria (22-24). However, the precise role of LrhA in *P. stewartii* and its relationship to both motility and RcsA is largely undefined.

To elucidate the downstream roles of RcsA and LrhA, phenotypic assays, RNA-Seq, qRT-PCR, GFP assays, and *in planta* virulence assays were used to evaluate the differences between the *rcsA* and *lrhA* deletion strains in comparison to the wild type. The identification of the most positively and negatively regulated targets for these two key downstream transcription factors provides greater insight into the coordinated regulation of genes in the QS network.

MATERIALS AND METHODS

Strains and growth conditions.

Strains and plasmids utilized in this study are listed in Table 2.1. *E. coli* strains were grown in Luria-Bertani (LB) (10 g/L tryptone, 5 g/L yeast extract, and 5 g/L NaCl) broth or on plates with 1.5% agar and *P. stewartii* strains were grown in either LB or Rich Minimal (RM) medium (1X M9 salts, 2% casamino acids, 1 mM $MgCl_2$, and 0.4% glucose). The growth medium was supplemented with nalidixic acid (Nal, 30 μg/ml), ampicillin (Ap, 100 μg/ml), kanamycin (Kn, 50 μg/ml), chloramphenicol (Cm, 30 μg/ml), or streptomycin (Str, 100 μg/ml) as required (see Table 2.1). *P. stewartii* strains were grown at 30°C, while *E. coli* strains were maintained at 37°C.

Construction of markerless deletion mutant strains.

Chromosomal deletions of *lrhA* and *rcsA* were constructed based on the Gateway system (Life Technologies, Grand Island, NY) and suicide vectors. Two 1kb fragments from upstream and downstream of the desired deletion region were first separately amplified using the primers designated in A.1 Table. After that, a two-step PCR reaction was performed in which the upstream and downstream segments for the specific gene were joined together and primers 1kbUPF-attB1 and 1kbDNR-attB2 (A.1 Table), specific to each construct, were added to

facilitate the Gateway BP reaction (Life Technologies). The final PCR products were cloned into the pGEM-T vector (Promega, Madison, WI) for sequencing before being transferred to the Gateway plasmids using BP Clonase II enzyme mix and then to the suicide vector pAUC40 (19) using LR Clonase II enzyme mix (Life Technologies). The resulting plasmids were transformed into competent *E. coli* DH5α *λ pir* cells (25). A tri-partite conjugation was used to transfer the suicide vector constructs into *P. stewartii* DC283 (26) using *E. coli* strain CC118*λpir* (27) carrying the conjugative helper plasmid pEVS104 (28) to facilitate suicide vector transfer. Selection for the first recombination event into the *P. stewartii* chromosome was carried out on LB agar supplemented with Nal and Str. Then, the recombinants were plated on LB (no salt) agar supplemented with 5% sucrose (29) to select for the desired double cross-over based on SacB activity. A screen for deletion strains was performed with colony PCR using a 3-primer reaction with primers for each gene corresponding to sites upstream, downstream, and inside the specific gene of interest (A.1 Table). DNA sequencing of appropriate PCR products was used to confirm the final deletion strains.

Construction of chromosomal complementation strains.

Complementation strains were constructed by generating a chromosomal insertion of the promoter and coding regions of the target gene into the neutral region downstream of *glmS* on the *P. stewartii* chromosome using the pUC18R6K-mini-Tn7-cat vector system developed by Choi *et al* (30). Specifically, primer sets with either *Eco*RI and *Xho*I or *Sac*I and *Spe*I sites (A.1 Table) were used to amplify the target regions and clone them into pGEM-T (Promega) for sequencing confirmation. DNA fragments of interest were moved into the pUC18R6K-mini-Tn7-cat vector using double digestion with *Eco*RI and *Xho*I or *Sac*I and *Spe*I (New England BioLabs (NEB), Ipswich, MA), followed by ligation with T4 DNA ligase (NEB) and

transformation into DH5α *λpir* or S17-1 *λpir* (31). These transformants served as the donor in the conjugation process with the appropriate deletion strain of *P. stewartii* as the recipient. Colony PCR reactions using one primer in the inserted gene and two primers flanking the *glmS* region (32) were conducted to screen for the presence of the chromosomal insertion. DNA sequencing of appropriate PCR products was used to validate the integrity of the complementation strains.

Phenotypic capsule production assay.

Wild-type, Δ*rcsA* and Δ*rcsA/rcsA*$^+$ strains were grown in LB supplemented with the appropriate antibiotics overnight at 30°C with shaking. The overnight cultures were used to inoculate fresh LB to an OD_{600} of 0.05 and grown at 30°C to OD_{600} of 0.2. The strains were then cross streaked on agar plates containing 0.1% casamino acids, 1% peptone, 1% glucose (CPG) and 1.5% agar (12). Capsule production was assessed qualitatively after 48 hours of incubation at 30°C by visually comparing the surface appearance of the streaks.

Phenotypic surface motility assay.

Swarming motility for the wild-type, Δ*lrhA,* and Δ*lrhA/lrhA*$^+$ strains was investigated under strict conditions to ensure a reproducible phenotype. Overnight cultures were diluted to an OD_{600} of 0.05 in LB broth and grown to an OD_{600} of 0.5. Five µl of cell culture at OD_{600} of 0.5 were spotted directly onto the agar surface of LB 0.4% agar quadrant plates supplemented with 0.4% glucose (21), which were poured on the same day of the experiment. Plates were put at room temperature for 30 min to 1 hour before incubating them lid-up in a closed box with a flat bottom inside the 30^{o}C incubator. Pictures of the plates were taken after 48 hours of incubation.

Transcriptome analysis methods.

The RNA-Seq method for analyzing the transcriptome of the wild-type *P. stewartii* DC283 strain and the two strains each carrying a deletion of one of the two genes, *lrhA* or *rcsA,*

has been previously published (15). Briefly, RNA was extracted from duplicate samples of each strain separately grown in RM medium using a Qiagen (Valencia, CA) miRNeasy RNA extraction kit. The total bacterial RNA was sent to the Virginia Bioinformatics Institute (VBI) (Virginia Tech, Blacksburg, VA) for Bioanalyzer quality analysis to insure RIN values greater than 9. The rRNA was depleted with an Epicentre (Madison, WI) Ribo-Zero Gram-negative Depletion Kit prior to Illumina (San Diego, CA) cDNA conversion and Illumina sequencing with single 50 bp reads.

RNA-Seq data analysis and qRT-PCR validation.

Data in the form of fastq files was received and aligned to the *P. stewartii* DC283 version 8 draft genome from NCBI (NZ_AHIE00000000.1) using the Geneious 7.0 software with the default "low sensitivity" settings, which also counted the numbers of reads aligning to each protein coding gene (excluding the highly repeated genes annotated as "transposase" or "IS66 ORF2 family protein"). Microsoft Excel was then used to compute an expression level for each gene by simply normalizing the read counts for each gene to the total number of mapped reads per million mapped reads (RPM) for each sample. Then the ratio of the RPM expression values (averaged over the two biological replicates) was used to compare the wild-type and deletion strain gene expression levels. (The raw read counts and RPM normalized expression levels for each of the four samples are available in the NCBI GEO database: Accession # GSE69064.) This analysis was used to select genes for qRT-PCR confirmation and validation of the RNA-Seq data. The validation genes were required to have (1) a greater than a four-fold change in average expression (RPM) between the two strains (wild-type and deletion), (2) greater than 100 reads mapped to the reference coding sequence in at least one of the samples and (3) reproducible levels of expression (< two-fold change) in duplicate trials. The Bioconductor R software

package “DESeq” (33) was also used to analyze the raw read counts using a more sophisticated gene expression normalization and error model to calculate multiple testing adjusted pvalues to estimate the statistical significance of detected gene expression changes. The fold changes (DESeq foldchange) determined by this second method were very similar to our Microsoft Excel analysis for the genes with 4-fold or greater change and the adjusted pvalues (DESeq padj) for those genes selected for qRT-PCR validation were all less than 0.012 (Tables 2.2 and 2.3).

The qRT-PCR method used for RNA-Seq validation has been previously described in Ramachandran et al (15). Briefly, each strain was grown in the same manner as for RNA-Seq. RNA was extracted using a miRNeasy RNA extraction kit (Qiagen) and converted to cDNA using the ABI High Capacity cDNA Reverse Transcription kit (Thermo Fisher Scientific, Waltham, MA). Primers (A.2 Table) designed using Primer Express software (ABI) were optimized and used to amplify ~100 bp regions of each gene of interest to determine the abundance of each transcript. The Pfaffl method was used to compare the wild type versus mutant abundance of a transcript to determine the fold regulation (34).

GFP fusion construction and testing.

A transcriptional fusion between the *rcsA* promoter and the gene for green fluorescent protein (GFP) was created using traditional molecular techniques. The *rcsA* promoter is located within the region -600 bp upstream of the annotated translation initiation codon (35). The restriction sites *Eco*RI and *Kpn*I were added to the 5’ and 3’ ends of the promoter sequence, respectively, through the PCR primers (A.1 Table). The PCR-amplified promoter fragment was ligated into pGEM-T (Promega) and sequenced. After restriction digestion of the pGEM-T construct and pPROBE’-GFP- [tagless] vector (36), a ligation produced the final pPROBE’-GFP-[tagless] vector containing the *rcsA* promoter. *E. coli* DH5α was transformed with this

plasmid construct which was then moved into wild-type *P. stewartii* DC283 via conjugation using a triparental mating with the pEVS104 helper plasmid. The conjugation plates were scanned using a Typhoon Trio Scanner (GE Healthcare, Pittsburgh, PA) set to use the blue laser to screen for GFP production. Subsequently, the Δ*lrhA* and Δ*lrhA/lrhA*$^{+}$ strains were conjugated to receive the same P_{rcsA} pPROBE'-GFP-[tagless] vector creating the desired reporter strains.

The transconjugates were grown in RM medium supplemented with Nal and Kn, overnight to an OD_{600} < 0.5 and then diluted in fresh RM to an OD_{600} of 0.025. The cultures were allowed to grow at 30°C with shaking at 250 RPM to an OD_{600} 0.5. The GFP production was monitored in 96-well plates using a Tecan Infinite 200 (Durham, NC) set with 485 excitation and 535 emission filters. The three strains were each analyzed in triplicate for one experiment, and the three wells were averaged together to establish the mean fluorescence. The average fluorescence reading from the blank was subtracted from the average fluorescence of each strain tested to remove background signal due to the medium. The normalized fluorescence was then divided by the OD_{600} for the sample to yield relative fluorescence readings/ OD_{600}. The three average relative fluorescence readings/ OD_{600} from three different experiments were averaged together and the overall standard error and two-tailed homoscedastic Student's t-test values calculated.

Plant virulence assay.

The procedure for conducting the virulence assays with *P. stewartii* strains in *Zea mays* seedlings was adapted from von Bodman et al. (12) with some modifications. Sweet corn seedlings (*Zea mays* cv. Jubilee, HPS Seed, Randolph, WI) were grown in Sunshine mix #1 soil in a growth chamber (Percival Scientific, Inc., Boone, IA) at 28°C, 80% relative humidity, 16 hours light and eight hours dark cycle, and at least 200 mE m^{-2} s^{-1} light intensity. Seedlings were

inoculated seven days after planting with five µl of bacterial culture grown to an OD_{600} of 0.2 in LB broth. Cells were washed and resuspended in an equal volume of phosphate buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4 and 2 mM KH_2PO_4, pH 7.4) prior to plant inoculation. An incision ~1 cm long was made ~1 cm above the soil line in the stem using a sterile needle (26 G 5/8, 15.9 mm, SUB-Q, Becton, Dickinson and Company, US). Then, the bacterial suspension was inoculated into the wound by moving the pipette tip across the wound five times. Fifteen germinated plants at day seven with two separate leaves and between 6-10 cm of height were inoculated for each bacterial strain tested. The plants were observed every other day after inoculation for up to 12 days post-infection to assess the virulence by two independent observers. Disease symptom severity was scored based on an arbitrary scale of five points, in which 0 = no symptoms; 1 = few scattered lesions; 2 = scattered water soaking symptoms; 3 = numerous lesions and slight wilting; 4 = moderately severe wilt; 5 = death. Both scores for each of the 15 plants under each treatment were averaged and then the data for each treatment were averaged together and used to calculate mean and standard error across the 15 plants.

Accession numbers.

The read data for the pairs of duplicate samples for the *P. stewartii* wild-type, Δ*rcsA*, and Δ*lrhA* strains, have been deposited in the NCBI Sequence Read Archive (SRA) with accession numbers, GSM1691841, GSM1691842, GSM1691843, GSM1691844, GSM1691845 and GSM1691846, respectively. An Excel file summarizing the differential gene expression in total counts and normalized reads per million (RPM), using the *P. stewartii* DC283 version 8 NCBI gene annotations, has been deposited in the NCBI Gene Expression Omnibus (GEO) database (GEO Accession GSE69064).

RESULTS

Deletion of *rcsA* or *lrhA* impacts phenotypic outputs.

Cross streaks of the *P. stewartii* wild-type, Δ*rcsA* and Δ*rcsA/rcsA*$^{+}$ strains on CPG agar demonstrated that deletion of *rcsA* yielded an easily visible decrease in the level of capsule production (Fig 2.1A). Chromosomal complementation of *rcsA* restored capsule synthesis to levels similar to the wild type (Fig 2.1A).

Separately, the *P. stewartii* wild-type, Δ*lrhA* and Δ*lrhA/lrhA*$^{+}$ strains were analyzed for motility. Under the test conditions employed, the wild-type strain exhibited unidirectional expansion the majority of the time as described by Herrera *et al* (21) (Fig 2.1B), but sometimes a more symmetrical expansion was observed. The Δ*lrhA* strain exhibited a different phenotype with a noticeably smaller occupied surface area (Fig 2.1B). Two chromosomally complemented strains of the Δ*lrhA* strain were constructed because a smaller promoter region (~600 bp upstream of *lrhA* coding region) strain could not complement the deletion (data not shown) whereas the strain with almost the entire intergenic region upstream of *lrhA* gene (~921 bp upstream) could (Fig 2.1B). This suggests that there are critical regulatory elements more than 600 bp upstream in the *lrhA* promoter region.

RNA-Seq analysis reveals the RcsA and LrhA regulons.

RNA-Seq data was acquired in duplicate for the *P. stewartii* wild-type, Δ*rcsA*, and Δ*lrhA* strains to analyze the global impact of the two regulators. Each strain yielded 17 to 19 million reads that were mapped to the protein coding genes on the *P. stewartii* DC283 genome (NZ_AHIE00000000.1). The normalized gene expression data (RPM) from the two trials was averaged and then the deletion strains were compared to the wild-type strain to determine the changes in gene expression when each transcription regulator was absent (Fig 2.2). Genes with

decreased expression in the deletion strains were considered to be positively regulated, either directly or indirectly, by RcsA or LrhA, respectively, in the wild-type strain. Conversely, genes with enhanced expression in the deletion strains were considered to be negatively regulated by the presence of LrhA or RcsA in the wild-type strain. A conservative four-fold change in RPM gene expression level was used as a selection criterion for genes to be further analyzed. An additional DESeq analysis supported the initial RPM data analysis. RcsA activates 11 genes and represses seven genes four-fold or greater (Table 2.2) and LrhA activates three genes and represses 23 genes four-fold or greater (Table 2.3). The most highly regulated RcsA-regulated genes are primarily related to capsule production, an activity previously shown to be under RcsA control (4). The most highly regulated genes in the LrhA regulon are mostly hypothetical proteins or proteins with just putative gene function. Interestingly, a 3.04-fold repression of *rcsA* by LrhA was detected during the RNA-Seq analysis hinting at a possible coordination of regulation between RcsA and LrhA.

qRT-PCR validates the RNA-Seq data.

Following the initial analysis of the RNA-Seq results, five target genes were selected from each of the putative RcsA and LrhA regulons to validate the RNA-Seq data via qRT-PCR. Some of the most highly regulated genes (Tables 2.2 and 2.3), which do not code for hypothetical proteins, were chosen for validation of the RNA-Seq data. The *rcsA* gene was also included in the analysis as its three-fold regulation by LrhA represented a potential feedback loop in the downstream QS system. The five genes tested for the RcsA regulon were: *wceG2, wza, argC,* CKS_3504, and CKS_2806. The five genes tested for the LrhA regulon were: CKS_3793, CKS_0458, CKS_5208, CKS_5211, and *rcsA.* The RNA-Seq data trends for the Δ*rcsA* and Δ*lrhA* strains were successfully validated by all five genes tested via qRT-PCR (Figs

2.3 and 2.4). Although there was variability in the absolute values of the fold changes between the RNA-Seq and qRT-PCR results, the regulation for a given gene shows similar trends (activation or repression) using both approaches. The qRT-PCR data also confirmed the initial observation from the RNA-Seq data that LrhA is repressing *rcsA* expression about three-fold. The interplay between these specific components of the *P. stewartii* QS system was thus further examined to build a more robust model of interactions.

RcsA is controlled by LrhA.

A *rcsA* promoter-GFP reporter transcription fusion was used to analyze expression of *rcsA* in *P. stewartii* DC283 wild-type, Δ*lrhA* and Δ*lrhA/lrhA*$^+$ strains (Fig 2.5). These experiments were performed at an OD_{600} of 0.5, the same growth conditions used for the RNA-Seq and qRT-PCR analysis. The relative fluorescence units (RFU) produced by the Δ*lrhA* strain were significantly more ($p< 0.05$) than either the wild-type strain or the Δ*lrhA/lrhA*$^+$ strain, suggesting LrhA represses *rcsA* either directly or indirectly. The results of the GFP assays further verified that LrhA negatively regulates *rcsA*, confirming the existence of a previously unrecognized feedback loop in the regulatory circuitry downstream of EsaR in the QS network; it is now recognized that both EsaR and LrhA negatively regulate *rcsA*.

RcsA and LrhA are involved in plant virulence.

The virulence of the Δ*rcsA* and Δ*lrhA* strains was compared to the wild type and respective complemented strains in a xylem-infection system. At day twelve post-infection, plants infected with the wild-type strain exhibit typical symptoms of the wilting stage and had an average disease severity score of ~3.5 while negative control plants infected with PBS were healthy with a score of ~0. The Δ*rcsA* strain also did not cause any severe disease symptoms with a score of ~0, while the Δ*lrhA* strain expressed an intermediate level of virulence with a

score of ~1.75 (Fig 2.6). The Δ*rcsA/rcsaA*$^{+}$ strain partially restored the virulence of *P. stewartii,* whereas the Δ*lrhA/lrhA*$^{+}$ strain fully complemented virulence. Both RcsA and LrhA clearly play roles in the pathogenicity of *P. stewartii.*

DISCUSSION

Previous studies demonstrated that the quorum-sensing master regulator EsaR directly represses *rcsA* and directly activates *lrhA* transcription (15, 35). As anticipated, phenotypic studies examining the impact of deletion of the genes encoding RcsA or LrhA resulted in noticeable effects on the production of capsule and the motility of *P. stewartii* cells, respectively. These important second-tier transcription factors downstream in the QS regulon were further examined for their role in global gene regulation via RNA-Seq analysis of gene expression in each of the two deletion strains compared with the wild-type strain. A qRT-PCR analysis confirmed that five genes for each deletion were regulated in the same manner shown by the RNA-Seq data. This allowed for increased confidence in the RNA-Seq data and the ability to draw more solid conclusions about the transcriptomes and the downstream network of gene regulation controlled by RcsA and LrhA.

Many of the most highly activated genes found in the RcsA regulon were related to capsule production: *wceG1, wceG2, wza, wzb, wzc, wceO, and wceF* (Table 2.2). Previous work has demonstrated that RcsA directly activates the promoters of the *cps* gene cluster in both *E. coli* and *P. stewartii* (4, 16). In *E. coli,* RcsA is one of two colonic acid capsular polysaccharide transcriptional activators, it is also self-activating, degraded by Lon-proteases, and requires RcsB to activate *cps* genes (16, 37). Similarly, in *P. stewartii* RcsA has been shown to activate genes required for EPS production as well as to self-activate its own gene (16, 35). The RNA-Seq data also revealed additional genes in the *P. stewartii* RcsA regulon that are directly or indirectly

suppressed by RcsA. Specifically, RcsA repressed several genes in the *cys* and *arg* operons more than four-fold (Table 2.2). The *cys* genes are involved in sulfate activation, which leads to cysteine biogenesis in *E. coli* (38). The genes *cysD* and *cysN* are in an operon and each encodes a subunit of the sulfate adenyltransferase complex (39, 40). The *arg* genes are involved in arginine biosynthesis in *E. coli* (41, 42). The three *arg* genes most regulated by RcsA are *argB* encoding acetylglutamate kinase, *argC* encoding N-acetyl-gamma-glutamylphosphate reductase, and *argI* encoding ornithine carbamoyltransferase. It is unclear why RcsA would repress the arginine and cysteine biosynthetic pathways and how this might relate to capsule production in *P. stewartii*.

P. stewartii LrhA is 77% identical at the amino acid level to its *E. coli* counterpart. In *E. coli*, LrhA is a LysR-type transcriptional factor negatively controlling motility, chemotaxis, flagellar biosynstheis (20) and type 1 fimbrial expression (43). However, unlike *E. coli*, where LrhA highly represses some genes related to flagellar function and chemotaxis by approximately 3-80 folds (20), the *P. stewartii* RNA-Seq data only indicated a ~two- to three-fold level of repression. On the other hand, LrhA does repress expression of two putative fimbrial subunits (encoded by CKS_0458 and CKS_0459) more than four-fold in *P. stewartii* (Table 2.3 and Fig 2.4), while the *E. coli* K12 strain MG1655 genome (CP009685.1) does not contain homologues of these genes. Thus, there are clear differences between the role of LrhA in *E. coli* and *P. stewartii*. Interestingly, the *P. stewartii* fimbrial subunit genes CKS_0458 and CKS_0459 were also previously found to be directly controlled by the master QS regulator EsaR (15). This is another example of coordinated control of gene expression at more than one level in the *P. stewartii* QS regulatory network.

Other *P. stewartii* annotated genes regulated by LrhA include CKS_5208, encoding rhamnosyltransferase I subunit B (RhlB), which is repressed. This enzyme may be involved in

surfactant production necessary to facilitate swarming motility. In addition, the gene most highly repressed by LrhA, CKS_5211, is located adjacent to CKS_5208 in the genome. Although originally annotated as a putative alpha/beta superfamily hydrolase/acyltransferase, CKS_5211 has very high homology with rhamnosyltransferase I subunit A (RhlA) of other sequenced strains of *Pantoea.* This further suggests an important role for LrhA in controlling surfactant production in *P. stewartii*. One annotated activated gene, CKS_3793, codes for cytochrome d ubiquinol oxidase subunit I, which presumably plays a role during aerobic respiration. The only other annotated gene found to be four-fold or more activated by LrhA was *wceO*, which is related to capsule synthesis and is also activated by RcsA, suggesting that its regulation is highly complex.

The RNA-Seq, qRT-PCR and transcriptional fusion experiments all confirmed the existence of a previously unknown feedback loop in the quorum-sensing network downstream of EsaR, with LrhA repressing RcsA. A developing model of this network suggests that there is tight coordinate control of key virulence factors in *P. stewartii* (Fig 2.7) (14, 15, 35) with the EsaR, RcsA, and LrhA proteins creating a coherent type three feed forward loop (44). The repressive role exerted by EsaR on *rcsA* transcription appears to be reinforced by the negative regulation of *rcsA* exerted by the EsaR-activated gene product LrhA. Negative control of *rcsA* expression by LrhA likely helps repress capsule production in the bacterium until the correct temporal point in disease progression. Positive autoregulation of RcsA in *P. stewartii* (35) will result in higher levels of the protein that escape proteolysis by Lon (16), thereby reinforcing the signal to increase capsule production at high cell density (35). Normally, QS tightly controls expression of *rcsA* so that the bacterium only produces high levels of capsule after migration to the xylem has occurred. If capsule is expressed too early, it can actually hinder the ability of *P.*

stewartii to cause disease (4). Alternatively, the Δ*rcsA* strain showed no sign of infection in the plant virulence assay. Thus, if the bacterium is incapable of producing capsule it also cannot form a biofilm in the xylem. The complement of *rcsA* however only partially restored the virulence of the deletion strain, perhaps indicating how important precise fine-tuning of *rcsA* expression is to disease outcome. The Δ*lrhA* strain significantly decreased the virulence of *P. stewartii*, but not to the degree seen with the Δ*rcsA* strain. This may in part be due to the removal of one level of regulation of *rcsA*, but is also probably due to impacts on the adhesion and motility of the bacterium, an area of on-going research. Since RcsA controls capsule production in *P. stewartii* and capsule is a main virulence factor for the bacterium, it would make sense for it to be regulated at multiple points in the regulatory network to ensure that the bacteria successfully migrate to the xylem before capsule and biofilm production begins in the corn plant. The tight control of RcsA by QS and the downstream network of transcriptional regulators, including RcsA itself (35) and LrhA, help ensure the precise timing of disease progression.

ACKNOWLEDGEMENTS

We thank Caroline Roper for graciously sharing plasmids and protocols, Revathy Ramachandran for help constructing the Δ*lrhA* strain, Tom Kuhar for providing corn seed and the laboratories of Rich Helm and Brenda Winkel for sharing their plant growth chamber.

REFERENCES

1. **Roper MC.** 2011. *Pantoea stewartii* subsp. *stewartii*: lessons learned from a xylem-dwelling pathogen of sweet corn. Molecular Plant Pathology **12:**628-637.
2. **Esker PD, Nutter FW.** 2002. Assessing the risk of Stewart's disease of corn through improved knowledge of the role of the corn flea beetle vector. Phytopathology **92:**668-670.
3. **Braun EJ.** 1982. Ultrastructural investigation of resistant and susceptible maize inbreds infected with *Erwinia stewartii*. Phytopathology **72:**159-166.
4. **Minogue TD, Carlier AL, Koutsoudis MD, von Bodman SB.** 2005. The cell density-dependent expression of stewartan exopolysaccharide in *Pantoea stewartii* ssp. *stewartii* is a function of EsaR-mediated repression of the *rcsA* gene. Molecular Microbiology **56:**189-203.
5. **Fuqua WC, Winans SC, Greenberg EP.** 1994. Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. Journal of Bacteriology **176:**269-275.
6. **Joint I, Allan Downie J, Williams P.** 2007. Bacterial conversations: talking, listening and eavesdropping. An introduction. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences **362:**1115-1117.
7. **von Bodman SB, Willey JM, Diggle SP.** 2008. Cell-cell communication in bacteria: united we stand. Journal of Bacteriology **190:**4377-4391.
8. **Whitehead NA, Barnard AM, Slater H, Simpson NJ, Salmond GP.** 2001. Quorum-sensing in Gram-negative bacteria. FEMS Microbiology Review **25:**365-404.
9. **de Kievit TR, Iglewski BH.** 2000. Bacterial quorum sensing in pathogenic relationships. Infection and Immunity **68:**4839-4849.
10. **Beck von Bodman S, Farrand SK.** 1995. Capsular polysaccharide biosynthesis and pathogenicity in *Erwinia stewartii* require induction by an N-acylhomoserine lactone autoinducer. Journal of Bacteriology **177:**5000-5008.
11. **von Bodman SB, Ball JK, Faini MA, Herrera CM, Minogue TD, Urbanowski ML, Stevens AM.** 2003. The quorum sensing negative regulators EsaR and $ExpR_{Ecc}$, homologues within the LuxR family, retain the ability to function as activators of transcription. Journal of Bacteriology **185:**7001-7007.
12. **von Bodman SB, Majerczak DR, Coplin DL.** 1998. A negative regulator mediates quorum-sensing control of exopolysaccharide production in *Pantoea stewartii* subsp. *stewartii*. Proceedings of the National Academy of Sciences of the United States of America **95:**7687-7692.
13. **Shong J, Huang YM, Bystroff C, Collins CH.** 2013. Directed evolution of the quorum-sensing regulator EsaR for increased signal sensitivity. ACS Chemical Biology **8:**789-795.
14. **Ramachandran R, Stevens AM.** 2013. Proteomic analysis of the quorum-sensing regulon in *Pantoea stewartii* and identification of direct targets of EsaR. Applied and Environmental Microbiology **79:**6244-6252.
15. **Ramachandran R, Burke AK, Cormier G, Jensen RV, Stevens AM.** 2014. Transcriptome-based analysis of the *Pantoea stewartii* quorum-sensing regulon and identification of EsaR direct targets. Applied and Environmental Microbiology **80:**5790-5800.
16. **Ebel W, Trempy JE.** 1999. *Escherichia coli* RcsA, a positive activator of colanic acid capsular polysaccharide synthesis, functions to activate its own expression. Journal of Bacteriology **181:**577-584.
17. **Keenleyside WJ, Jayaratne P, MacLachlan PR, Whitfield C.** 1992. The *rcsA* gene of *Escherichia coli* O9:K30:H12 is involved in the expression of the serotype-specific group I K (capsular) antigen. Journal of Bacteriology **174:**8-16.

18. **Virlogeux I, Waxin H, Ecobichon C, Lee JO, Popoff MY.** 1996. Characterization of the *rcsA* and *rcsB* genes from *Salmonella typhi: rcsB* through *tviA* is involved in regulation of Vi antigen synthesis. Journal of Bacteriology **178:**1691-1698.
19. **Carlier A, Burbank L, von Bodman SB.** 2009. Identification and characterization of three novel EsaI/EsaR quorum-sensing controlled stewartan exopolysaccharide biosynthetic genes in *Pantoea stewartii* ssp. *stewartii*. Molecular Microbiology **74:**903-913.
20. **Lehnen D, Blumer C, Polen T, Wackwitz B, Wendisch VF, Unden G.** 2002. LrhA as a new transcriptional key regulator of flagella, motility and chemotaxis genes in *Escherichia coli*. Molecular Microbiology **45:**521-532.
21. **Herrera CM, Koutsoudis MD, Wang X, von Bodman SB.** 2008. *Pantoea stewartii* subsp. *stewartii* exhibits surface motility, which is a critical aspect of Stewart's wilt disease development on maize. Molecular Plant-Microbe Interactions **21:**1359-1370.
22. **Francez-Charlot A, Laugel B, Van Gemert A, Dubarry N, Wiorowski F, Castanie-Cornet MP, Gutierrez C, Cam K.** 2003. RcsCDB His-Asp phosphorelay system negatively regulates the *flhDC* operon in *Escherichia coli*. Molecular Microbiology **49:**823-832.
23. **Clemmer KM, Rather PN.** 2007. Regulation of *flhDC* expression in *Proteus mirabilis*. Research in Microbiology **158:**295-302.
24. **Wang Q, Zhao Y, McClelland M, Harshey RM.** 2007. The RcsCDB signaling system and swarming motility in *Salmonella enterica* serovar *typhimurium*: dual regulation of flagellar and SPI-2 virulence genes. Journal of Bacteriology **189:**8447-8457.
25. **Kvitko BH, Bruckbauer S, Prucha J, McMillan I, Breland EJ, Lehman S, Mladinich K, Choi KH, Karkhoff-Schweizer R, Schweizer HP.** 2012. A simple method for construction of *pir*+ Enterobacterial hosts for maintenance of R6K replicon plasmids. BMC Research Notes **5:**157.
26. **Dolph PJ, Majerczak DR, Coplin DL.** 1988. Characterization of a gene cluster for exopolysaccharide biosynthesis and virulence in *Erwinia stewartii*. Journal of Bacteriology **170:**865-871.
27. **Herrero M, de Lorenzo V, Timmis KN.** 1990. Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in gram-negative bacteria. Journal of Bacteriology **172:**6557-6567.
28. **Stabb EV, Ruby EG.** 2002. RP4-based plasmids for conjugation between *Escherichia coli* and members of the *Vibrionaceae*. Methods in Enzymology **358:**413-426.
29. **Kaniga K, Delor I, Cornelis GR.** 1991. A wide-host-range suicide vector for improving reverse genetics in gram-negative bacteria: inactivation of the *blaA* gene of *Yersinia enterocolitica*. Gene **109:**137-141.
30. **Choi KH, Gaynor JB, White KG, Lopez C, Bosio CM, Karkhoff-Schweizer RR, Schweizer HP.** 2005. A Tn7-based broad-range bacterial cloning and expression system. Nature Methods **2:**443-448.
31. **Labes M, Puhler A, Simon R.** 1990. A new family of RSF1010-derived expression and *lac*-fusion broad-host-range vectors for gram-negative bacteria. Gene **89:**37-46.
32. **Burbank L, Roper MC.** 2014. OxyR and SoxR modulate the inducible oxidative stress response and are implicated during different stages of infection for the bacterial phytopathogen *Pantoea stewartii* subsp. *stewartii*. Molecular Plant-Microbe Interactions **27:**479-490.
33. **Anders S, Huber W.** 2010. Differential expression analysis for sequence count data. Genome Biology **11:**R106.
34. **Pfaffl MW.** 2001. A new mathematical model for relative quantification in real-time RT–PCR. Nucleic Acids Research **29:**e45.
35. **Carlier AL, von Bodman SB.** 2006. The *rcsA* promoter of *Pantoea stewartii* subsp. *stewartii* features a low-level constitutive promoter and an EsaR quorum-sensing-regulated promoter. Journal of Bacteriology **188:**4581-4584.

36. **Miller WG, Leveau JH, Lindow SE.** 2000. Improved *gfp* and *inaZ* broad-host-range promoter-probe vectors. Molecular Plant-Microbe Interactions **13:**1243-1250.
37. **Hagiwara D, Sugiura M, Oshima T, Mori H, Aiba H, Yamashino T, Mizuno T.** 2003. Genome-wide analyses revealing a signaling network of the RcsC-YojN-RcsB phosphorelay system in *Escherichia coli*. Journal of Bacteriology **185:**5735-5746.
38. **Awano N, Wada M, Mori H, Nakamori S, Takagi H.** 2005. Identification and functional analysis of *Escherichia coli* cysteine desulfhydrases. Applied and Environmental Microbiology **71:**4149-4152.
39. **Malo MS, Loughlin RE.** 1990. Promoter elements and regulation of expression of the *cysD* gene of *Escherichia coli* K-12. Gene **87:**127-131.
40. **Leyh TS, Taylor JC, Markham GD.** 1988. The sulfate activation locus of *Escherichia coli* K12: cloning, genetic, and enzymatic characterization. The Journal of Biological Chemistry **263:**2409-2416.
41. **Ludovice M, Martin JF, Carrachas P, Liras P.** 1992. Characterization of the *Streptomyces clavuligerus argC* gene encoding N-acetylglutamyl-phosphate reductase: expression in *Streptomyces lividan*s and effect on clavulanic acid production. Journal of Bacteriology **174:**4606-4613.
42. **Parsot C, Boyen A, Cohen GN, Glansdorff N.** 1988. Nucleotide sequence of *Escherichia coli argB* and *argC* genes: comparison of N-acetylglutamate kinase and N-acetylglutamate-gamma-semialdehyde dehydrogenase with homologous and analogous enzymes. Gene **68:**275-283.
43. **Blumer C, Kleefeld A, Lehnen D, Heintz M, Dobrindt U, Nagy G, Michaelis K, Emody L, Polen T, Rachel R, Wendisch VF, Unden G.** 2005. Regulation of type 1 fimbriae synthesis and biofilm formation by the transcriptional regulator LrhA of *Escherichia coli*. Microbiology **151:**3287-3298.
44. **Alon U.** 2007. Network motifs: theory and experimental approaches. Nature Review Genetics **8:**450-461.
45. **Grant SG, Jessee J, Bloom FR, Hanahan D.** 1990. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. Proceedings of the National Academy of Sciences of the United States of America **87:**4645-4649.
46. **Raleigh EA, Murray NE, Revel H, Blumenthal RM, Westaway D, Reith AD, Rigby PW, Elhai J, Hanahan D.** 1988. McrA and McrB restriction phenotypes of some *E. coli* strains and implications for gene cloning. Nucleic Acids Research **16:**1563-1575.

Table 2.1: Strains and plasmids used in the study

Strains	Genotype and notes[a]	References
***Pantoea stewartii* strains**		
DC283	Wild-type strain; Nal^r	(26)
Δ*lrhA*	Unmarked deletion of *lrhA* coding sequence; Nal^r	This study
Δ*lrhA/lrhA*$^+$	DC283 Δ*lrhA* with chromosomal complementation of *lrhA* and its promoter downstream of *glmS*; Nal^r Cm^r	This study
Δ*rcsA*	Unmarked deletion of *rcsA* coding sequence; Nal^r	This study
Δ*rcsA/rcsA*$^+$	DC283 Δ*rcsA* with chromosomal complementation of *rcsA* and its promoter downstream of *glmS*; Nal^r Cm^r	This study
***Escherichia coli* strains**		
Top 10	F$^-$ *mcrA* Δ(*mrr-hsdRMS-mcrBC*) *Φ80dlacZΔM15* Δ*lac*X74 *deoR recAI araD139* Δ(*ara-leu)7697 galU galK rpsL (*Str^r*) endA1 nupG*	(45)
DH5α	*F- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(rK- mK+)*	(46)
DH5α *λpir*	*F- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(rK- mK+), λpir*	(25)
CC118 *λpir*	*Δ(ara-leu), araD,* Δ*lacX74, galE, galK, phoA20, thi-1, rpsE, rpoB, argE(Am), recA1, λpir*	(27)
S17-1 *λpir*	*recA pro hsdR RP4-2-Tc::Mu-Km::Tn7*	(31)
Plasmids		
pGEM-T	Cloning vector, Ap^r	Promega
pDONR201	Entry vector in the Gateway system, Kn^r	Life Technologies
pAUC40	Suicide vector pKNG101::*attR-ccdB*-Cm^R; Cm^r, Str^r, *sacB*	(19)
pEVS104	Conjugative helper plasmid, *tra trb;* Kn^r	(28)
pUC18R6K-mini-Tn7-cat	Tn7 vector for chromosomal integration into the intergenic region downstream of *glmS*; Cm^r, Ap^r	(30)
pPROBE'GFP[tagless] P_{rcsA}	pPROBE'GFP[tagless] vector with the promoter of *rcsA;* Kn^r	This study

[a] Ap^r, ampicillin resistance; Nal^r, nalidixic acid resistance; Kn^r, kanamacyin resistance; Gm^r, gentamycin resistance; Cm^r, chloramphenicol resistance; Str^r, streptomycin resistance

Table 2.2: List of genes differentially expressed 4-fold or more in the Δ*rcsA* RNA-Seq data

Accession #	Locus_tag	GeneID	Product	RPM Fold Change	DESeq Fold Change	DESeq padj
Activated by RcsA						
ACV-0288878	CKS_4672		secreted protein	17.69	16.08	1.2E-01
ACV-0289544	CKS_2241	*wceG1*	undecaprenyl-phosphate UDP-galactose phosphotransferase	12.60	11.89	2.1E-15
ACV-0290198	CKS_2799	*osmB*	lipoprotein	9.12	8.08	1.0E+00
ACV-0286879	CKS_2708	*wceG2*	undecaprenyl-phosphate UDP-galactose phosphotransferase	7.78	7.47	2.7E-08
ACV-0288877	CKS_4671		putative outer membrane lipoprotein	7.44	6.99	5.2E-09
ACV-0289541	CKS_2244	*wza*	polysaccharide export protein	6.26	5.75	1.2E-02
ACV-0288876	CKS_4670		YmcB family protein	5.79	5.53	3.0E-07
ACV-0289540	CKS_2245	*wzb*	phosphotyrosine-protein phosphatase	5.16	4.74	6.2E-02
ACV-0289539	CKS_2246	*wzc*	tyrosine-protein kinase	5.13	4.70	2.8E-02
ACV-0288191	CKS_4022	*wceO*	beta-16-glucosyltransferase	4.75	4.46	6.3E-06
ACV-0289534	CKS_2251	*wceF*	exopolysaccharide biosynthesis protein	4.11	3.78	9.7E-02
Repressed by RcsA						
ACV-0290191	CKS_2806		putative formate dehydrogenase oxidoreductase protein	50.38	54.64	5.7E-34
ACV-0289299	CKS_3504		cytosine/purine/uracil/thiamine/allantoin permease family protein	8.85	9.44	1.3E-11
ACV-0289953	CKS_1065	*argC*	N-acetyl-gamma-glutamylphosphate reductase	5.25	5.50	8.6E-08
ACV-0289954	CKS_1064	*argB*	acetylglutamate kinase	5.14	5.46	2.3E-07
ACV-0291072	CKS_1283	*cysD*	sulfate adenylyltransferase subunit 2	5.07	5.57	3.0E-07
ACV-0291071	CKS_1282	*cysN*	sulfate adenylyltransferase subunit 1	4.29	4.64	4.7E-06
ACV-0289145	CKS_4942	*argI*	ornithine carbamoyltransferase 1	4.20	4.54	1.1E-05

Table 2.3: List of genes differentially expressed 4-fold or more in the Δ*lrhA* RNA-Seq data

Accession #	Locus_tag	GeneID	Product	RPM Fold Change	DESeq Fold Change	DESeq padj
Activated by LrhA						
ACV-0289574	CKS_2211		hypothetical protein	7.45	6.97	7.9E-04
ACV-0288191	CKS_4022	*wceO*	beta-16-glucosyltransferase	7.43	6.80	3.7E-11
ACV-0291275	CKS_3793		cytochrome d ubiquinol oxidase subunit I	4.72	4.30	3.4E-05
Repressed by LrhA						
ACV-0287751	CKS_5211		putative alpha/beta superfamily hydrolase/acyltransferase	58.54	61.87	8.5E-13
ACV-0287748	CKS_5208		rhamnosyltransferase I subunit B	17.02	18.57	2.0E-05
ACV-0290189	CKS_2808		hypothetical protein	15.00	16.26	3.5E-16
ACV-0285926	CKS_2106		hypothetical protein	11.06	11.98	5.4E-17
ACV-0285999	CKS_2612		phage holin	8.80	9.15	3.5E-14
ACV-0290526	CKS_0458		putative fimbrial subunit	8.39	9.32	1.1E-09
ACV-0286015	CKS_2628		hypothetical protein	6.69	7.14	2.9E-09
ACV-0286005	CKS_2618		hypothetical protein	6.46	6.83	3.4E-04
ACV-0286003	CKS_2616		phage protein	5.73	5.92	5.4E-04
ACV-0286012	CKS_2625		hypothetical protein	5.58	5.79	9.5E-03
ACV-0286019	CKS_2632		hypothetical protein	5.53	5.67	3.3E-02
ACV-0286011	CKS_2624		hypothetical protein	5.51	5.77	1.3E-07
ACV-0286029	CKS_2642		hypothetical protein	5.46	5.63	1.1E-04
ACV-0286016	CKS_2629		phage tail sheath protein FI	5.14	5.49	6.4E-09
ACV-0290525	CKS_0459		putative fimbrial subunit	5.02	5.57	2.1E-08
ACV-0290434	CKS_0551		carbonic anhydrase	4.84	5.25	5.9E-08
ACV-0286009	CKS_2622		hypothetical protein	4.77	5.17	3.6E-03
ACV-0286013	CKS_2626		hypothetical protein	4.29	4.60	3.4E-05
ACV-0286031	CKS_2644		hypothetical protein	4.29	4.57	3.5E-04
ACV-0286018	CKS_2631		hypothetical protein	4.27	4.44	1.4E-01
ACV-0286014	CKS_2627		hypothetical protein	4.26	4.47	7.7E-03
ACV-0286028	CKS_2641		phage baseplate assembly protein V	4.17	4.55	1.4E-02
ACV-0286022	CKS_2635		hypothetical protein	4.00	4.14	3.4E-05

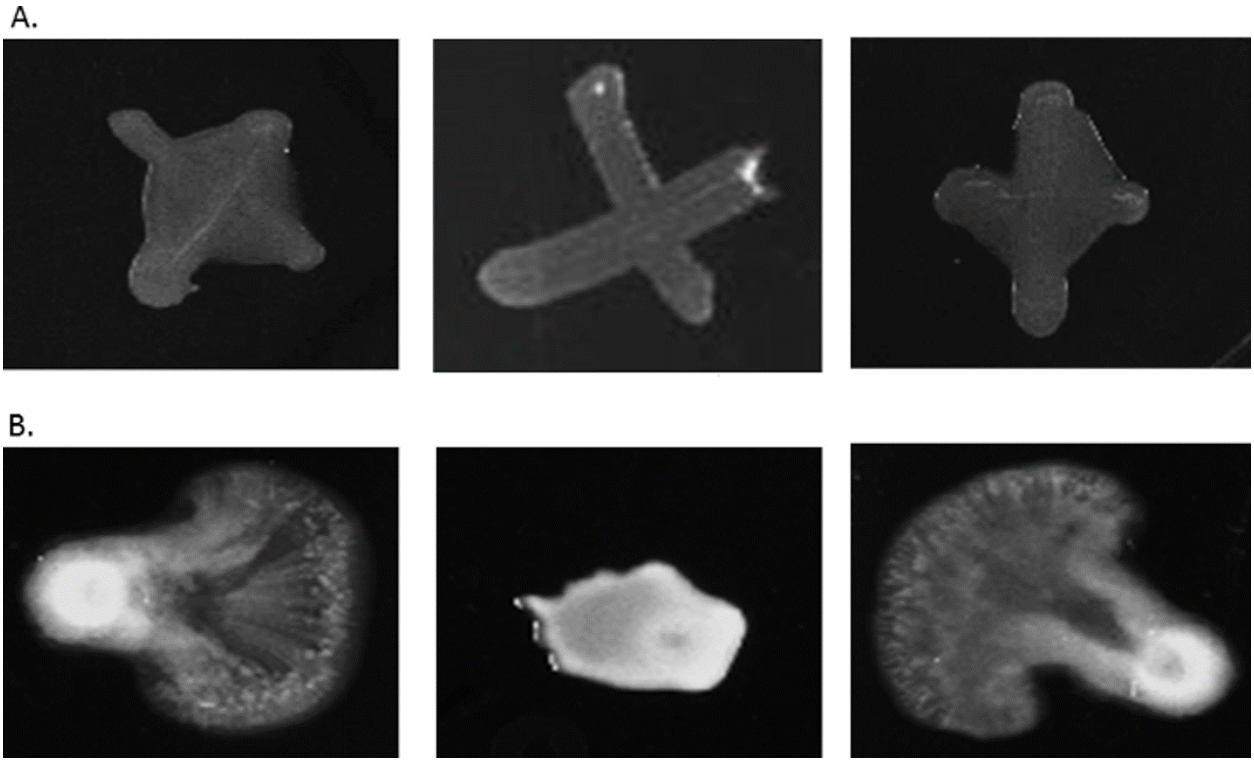

Fig 2.1. Impact of RcsA and LrhA on phenotype of *P. stewartii.* Panel A shows an analysis of capsule production in *P. stewartti* DC283 wild-type, Δ*rcsA* mutant and Δ*rcsA/rcsA*$^{+}$ complementation strains (left to right). Differences in capsule production are apparent in the regions between the arms of the X-cross streak. Panel B shows an analysis of swarming motility in wild-type, Δ*lrhA* mutant and Δ*lrhA/lrhA*$^{+}$ complementation strains (left to right). All pictures for panel A or B, respectively, were taken at the same magnification after 48 hours of incubation.

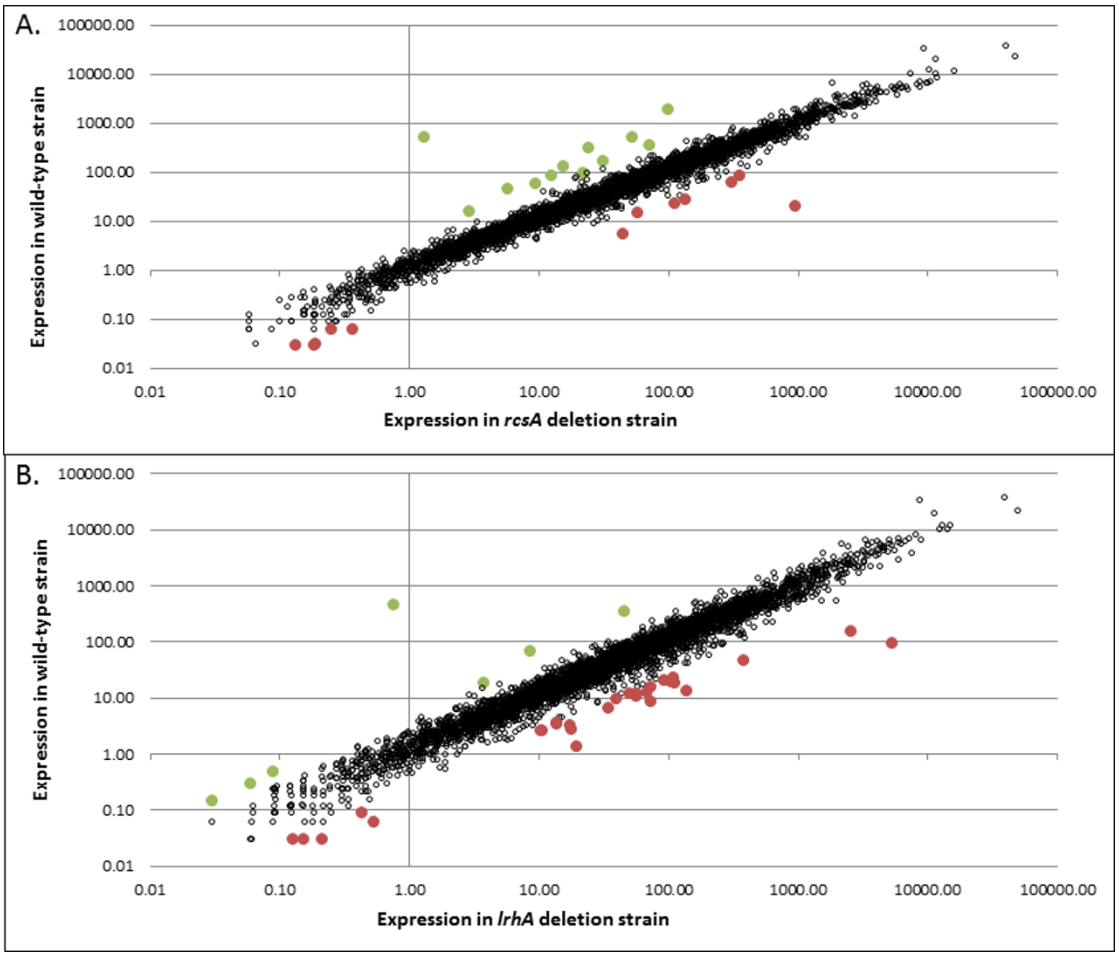

Fig 2.2. Differential mRNA expression during QS. Whole transcriptome data of the *P. stewartii* DC283 wild-type strain compared to the Δ*rcsA* strain (panel A) or the Δ*lrhA* strain (panel B). An open circle is used to represent each gene. Those filled with green are activated (levels of expression are four-fold or lower in the deletion strain) and those filled with red are repressed (levels of expression are four-fold or higher in the deletion strain), by either RcsA (panel A) or LrhA (panel B). The two extreme outlier points, with RPM expression >100 in the wild-type and ~1 in the deletion strains, represent the deleted genes *rcsA* or *lrhA*, respectively. All of the other colored points with RPM expression >1 in both samples are tabulated in Tables 2 and 3. The RPM change of normalized expression for most genes fall tightly around a line of slope of 1, indicating that they are approximately equally expressed in both strains.

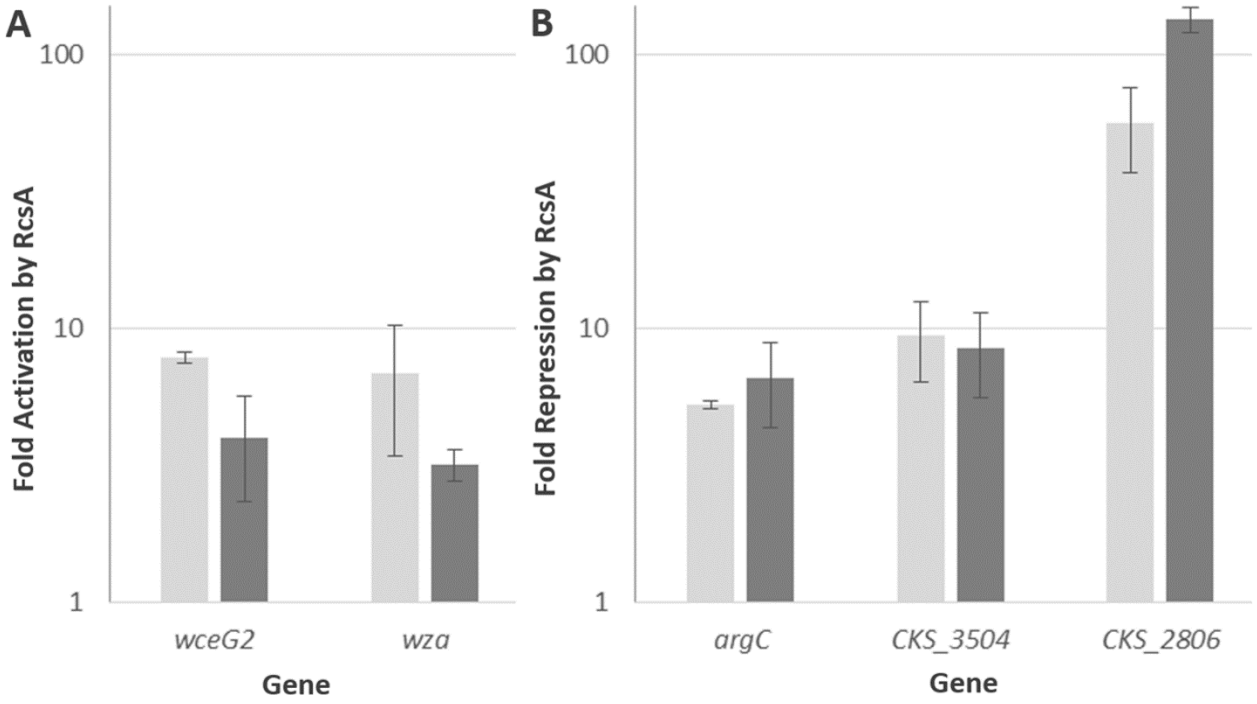

Fig 2.3. Validating transcriptional control of select genes by RcsA. Changes in gene expression were compared between the RNA-Seq analysis (light grey) and qRT-PCR assays (dark grey) for five genes regulated by RcsA. Y-axis represents the fold activation (panel A) or repression (panel B) on a logarithmic scale in the presence of RcsA. RNA-Seq results are averages of two experimental samples and qRT-PCR data represent two experimental samples analyzed in triplicate. Error bars were estimated using the sample standard error of the fold-change across the two independent biological replicates for both RNA-Seq and qRT-PCR.

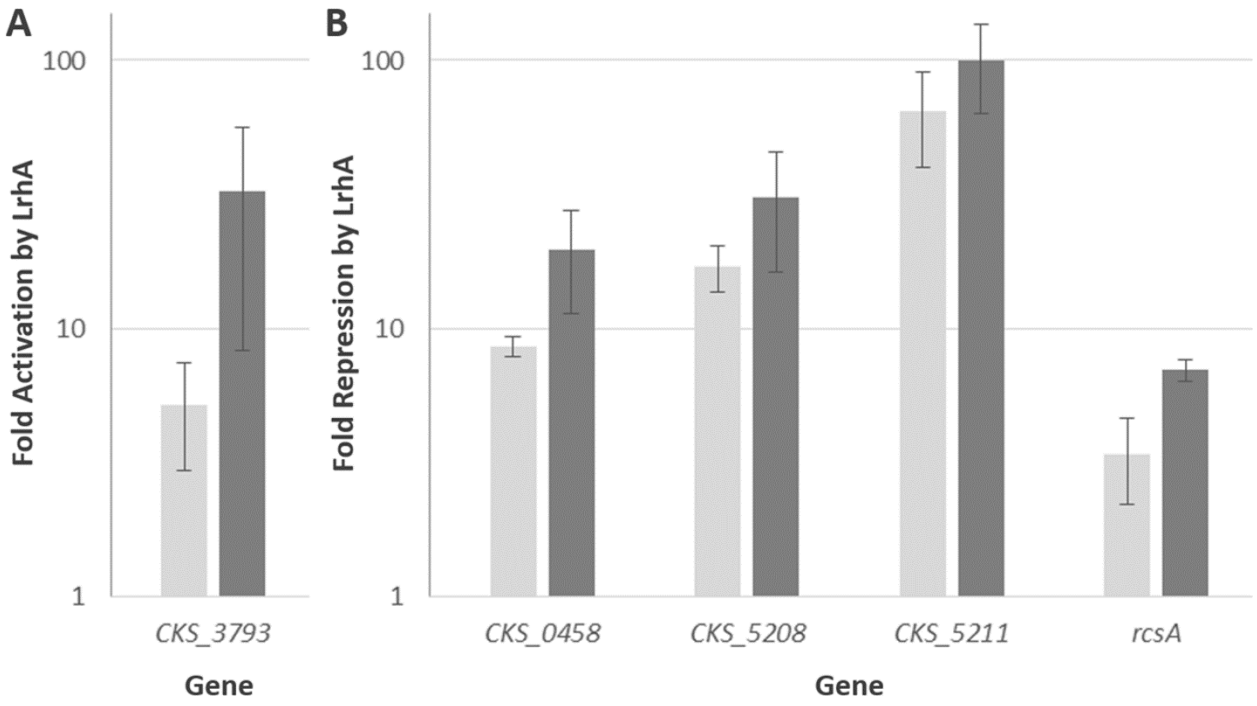

Fig 2.4. Validating transcriptional control of select genes by LrhA. Changes in gene expression were compared between the RNA-Seq analysis (light grey) and qRT-PCR assays (dark grey) for five genes regulated by LrhA. Y-axis represents the fold activation (panel A) or repression (panel B) on a logarithmic scale in the presence of LrhA. RNA-Seq results are averages of two experimental samples and qRT-PCR data represent two experimental samples analyzed in triplicate. Error bars were estimated using the sample standard error of the fold-change across the two independent biological replicates for both RNA-Seq and qRT-PCR.

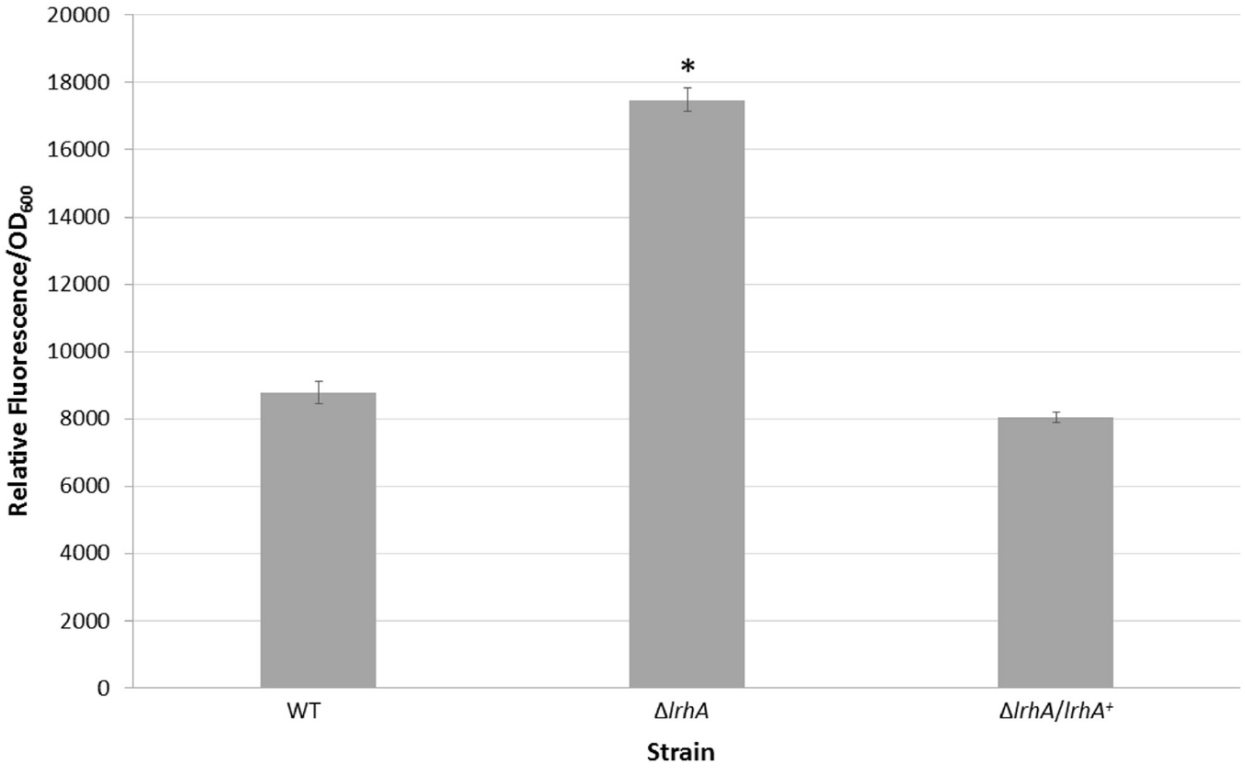

Fig 2.5. Expression from the *rcsA* promoter. A GFP reporter was used to measure levels of transcription from the *rcsA* promoter. Strains were grown to an OD_{600} of 0.5 and the average fluorescence/OD_{600} was measured. As indicated by an asterisk, expression from the Δ*lrhA* strain is significantly higher ($p < 0.05$) than either the wild-type strain or the Δ*lrhA/lrhA*$^+$ strain using a two-tailed homoscedastic Student's t-test. This indicates that LrhA normally represses expression of *rcsA* in the wild-type strain. Data represents three experimental samples analyzed in triplicate. Error bars denote standard error.

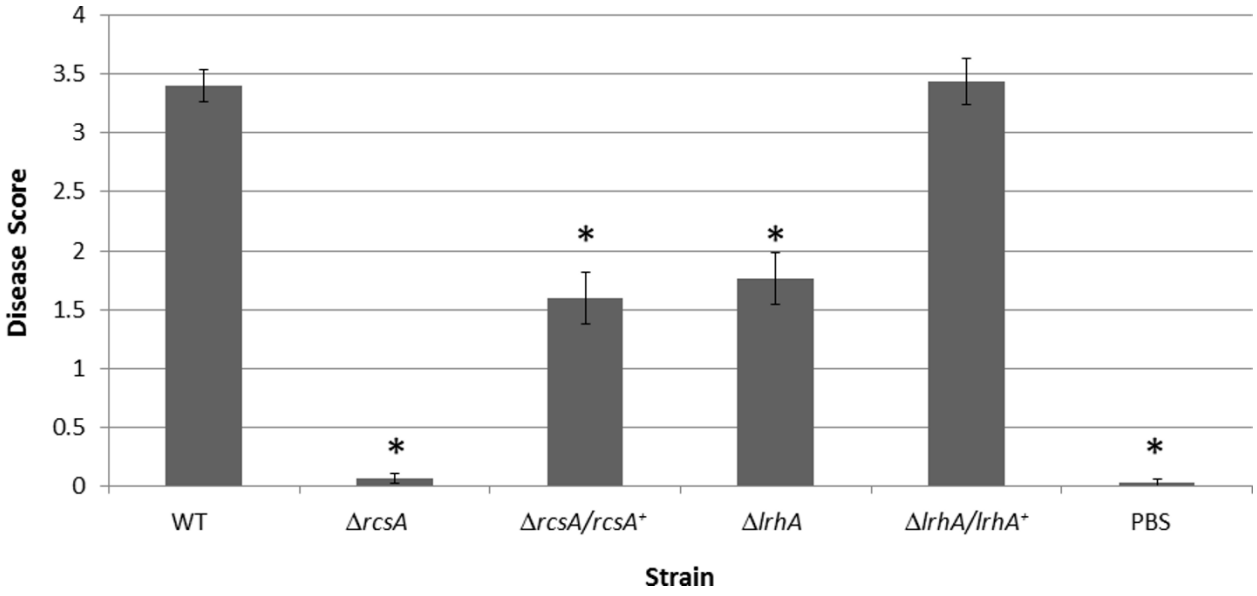

Fig 2.6. Plant assays testing the role of RcsA or LrhA in virulence. Data shown is the average score of disease for Day 12 of an infection assay performed with 15 plants inoculated with *P. stewartii* DC283 strains: wild type (WT), Δ*rcsA*, Δ*rcsA/rcsA*$^{+}$, Δ*lrhA*, Δ*lrhA/lrhA*$^{+}$, or PBS as a negative control. The asterisks (*) represent strains that are statistically significantly different ($p < 0.05$) from the wild-type strain using a two-tailed homoscedastic Student's t-test. Error bars denote standard error.

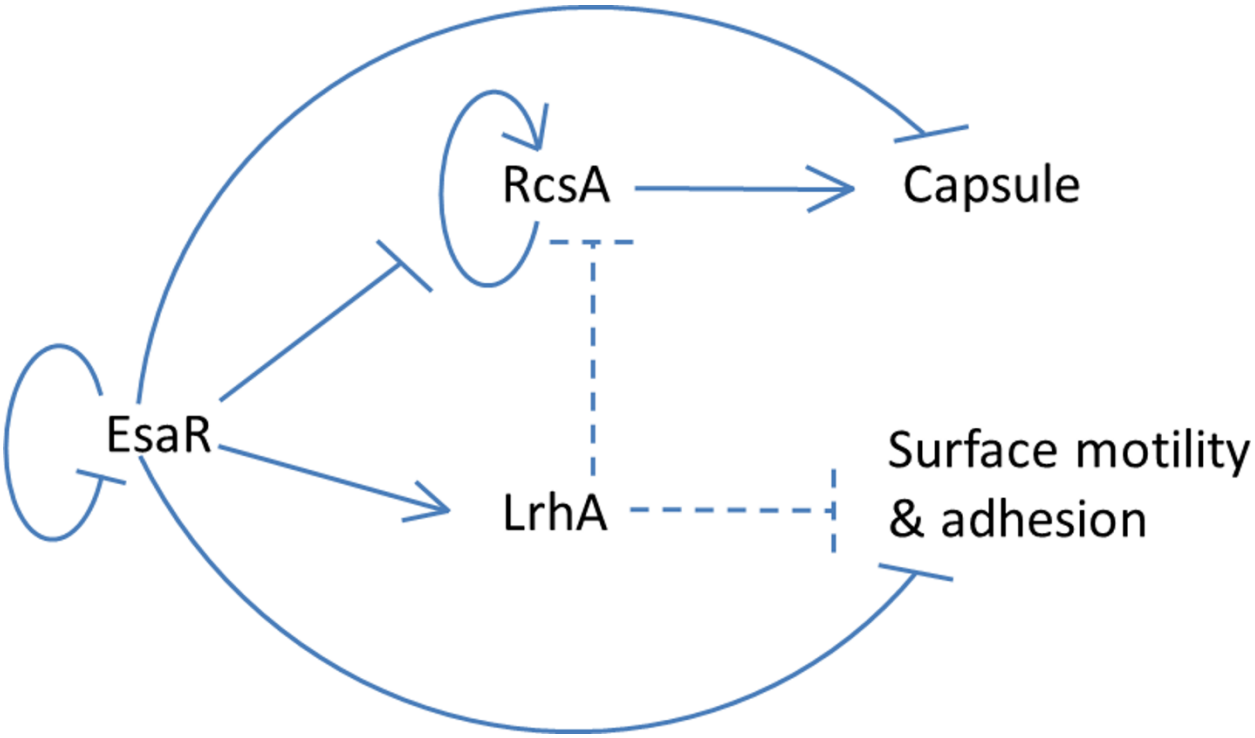

Fig 2.7. Model of the quorum-sensing regulatory network in *P. stewartii*. See the text for details. Solid lines indicate known direct regulatory control. Dashed lines indicate either direct or indirect control found in the present study. Arrows represent activation and T lines represent repression.

Chapter Three

PART I:

Complete Genome Assembly of *Pantoea stewartii* subsp. *stewartii* DC283, a Corn Pathogen

Duy An Duong, Ann M. Stevens, Roderick V. Jensen[#]. 2017. Complete genome assembly of *Pantoea stewartii* subsp. *stewartii* DC283, a corn pathogen. Genome Announcements 5: e00435-17. https://doi.org/10.1128/genomeA.00435-17.

[#]Address correspondence to Roderick V. Jensen, rvjensen@vt.edu

Running Head: Complete Genome Assembly of *Pantoea stewartii*

PART II:

Expanded Analysis of the *Pantoea stewartii* subsp. *stewartii* DC283 Complete Genome Reveals Plasmid-borne Virulence Factors

Duy An Duong, Ann M. Stevens, Roderick V. Jensen

PART I: Complete Genome Assembly of *Pantoea stewartii* subsp. *stewartii* DC283, a Corn Pathogen

ABSTRACT

The phytopathogen *Pantoea stewartii* subsp. *stewartii* DC283 causes Stewart's wilt disease in corn after transmission from the corn flea beetle insect vector. Here we report that the complete, annotated genome of *P. stewartii* DC283 has been fully assembled into one circular chromosome, ten circular plasmids and one linear phage.

MAIN TEXT

Pantoea stewartii subsp. *stewartii* (*P. stewartii*) is a Gram negative, gamma-proteobacterium, native to North America, which causes Stewart's wilt disease in corn resulting in economic loss and quarantine of exports. Transmission to plants primarily occurs through seeds or inoculation from the corn flea beetle (1, 2). *P. stewartii* DC283 is the wild-type reference strain used to study pathogenesis (3-5). It is a nalidixic acid-resistant mutant of the original 1976 isolate, SS104, from *Zea mays* (6, 7). In 2012 a draft assembly of the genome with 65 contigs (NCBI GenBank: AHIE00000000.1) was released. However, the completion of this genome has been complicated by large numbers of repetitive transposon sequences often spanning 1,500 bp. Using Illumina Mate Pair sequencing with 3,500 bp insert size, this study reports the complete genome assembly of *P. stewartii* DC283.

Genomic DNA was extracted from an overnight culture of *P. stewartii* DC283 grown in Luria Bertani medium using a Qiagen DNeasy Blood & Tissue Kit per the manufacturer's recommendations for Gram-negative bacteria. On-column RNase treatment was applied before DNA elution to yield high-quality DNA for Nextera Mate Pair preparation and Illumina MiSeq

sequencing using 250 bp Mate Pair Reads with 3,500 bp inserts. The mate-pair reads were first aligned to the 65 contigs of the reference genome using the Geneious V9.1.2 software package (Biomatters Ltd.) 'Map to Reference' function. Unmapped mate-pair reads were then *de novo* assembled to identify missing sequences and rearrangements of the reference contigs. Finally, the end-sequences of all of the remaining contigs were extended ~3,500 bp by *de novo* assembling mate-pair reads where one mate maps to the last 3,500 bp of the contig. This procedure was sufficient to link most of the reference contigs, leading to the complete assembly of the ten circular plasmids and of the main chromosome into three large segments. The chromosome segments were connected into a closed, circular sequence using long-range PCR (Qiagen LongRange PCR kit). In addition, a linear phage was identified with similarity to the N15 phage-plasmid in *Escherichia coli* (8). This resulted in ~200X coverage of the genomic sequence and between ~200X to ~3,000X coverage for the plasmids.

Automatic annotation for the *P. stewartii* DC283 genome was performed using the NCBI prokaryotic genome annotation pipeline (9, 10). The whole genome consists of 5,314,092 bp (53.8% G+C contents) with 5,625 coding sequences, 21 rRNAs and 73 tRNAs. The chromosome is 4,528,215 bp, the plasmids range in size from 4,277-304,641 bp and the linear phage is 47,186 bp. More than 460 sequences encoding repetitive transposases were found in the complete genome. Interestingly, the two type III secretion systems that play important roles in colonization of insect and plant hosts (11) were found to be located on two separated mega-plasmids. In addition, a 66-kb region was newly assembled in this genome compared to the reference genome.

Nucleotide sequence accession number. The annotated genome assembly of *P. stewartii* DC283 is available in GenBank under accession no. CP017581-CP017592.

ACKNOWLEDGMENTS

We thank Susanne von Bodman for sharing her stock strain of *P. stewartii* DC283 with us. This work was funded by the Fralin Life Science Institute and Biological Sciences Department, Virginia Tech. Mate-pair library construction and Illumina sequencing were conducted by Genomics Research laboratory at the Biocomplexity Institute at Virginia Tech.

REFERENCES

1. **Von Bodman SB, Bauer WD, Coplin DL.** 2003. Quorum sensing in plant-pathogenic bacteria. Annual Review of Phytopathology **41:**455-482.
2. **Roper MC.** 2011. *Pantoea stewartii* subsp. *stewartii*: lessons learned from a xylem-dwelling pathogen of sweet corn. Molecular Plant Pathology **12:**628-637.
3. **Ham JH, Majerczak D, Ewert S, Sreerekha MV, Mackey D, Coplin D.** 2008. WtsE, an AvrE-family type III effector protein of *Pantoea stewartii* subsp. *stewartii*, causes cell death in non-host plants. Molecular Plant Pathology **9:**633-643.
4. **Burbank L, Roper MC.** 2014. OxyR and SoxR modulate the inducible oxidative stress response and are implicated during different stages of infection for the bacterial phytopathogen *Pantoea stewartii* subsp. *stewartii*. Molecular Plant-Microbe Interactions **27:**479-490.
5. **Kernell Burke A, Duong DA, Jensen RV, Stevens AM.** 2015. Analyzing the transcriptomes of two quorum-sensing controlled transcription factors, RcsA and LrhA, important for *Pantoea stewartii* virulence. PLOS ONE **10:**e0145358.
6. **Coplin DL, Rowan RG, Chisholm DA, Whitmoyer RE.** 1981. Characterization of plasmids in *Erwinia stewartii*. Applied and Environmental Microbiology **42:**599-604.
7. **Coplin DL, Majerczak DR, Zhang Y, Kim W, Jock S, Geider K.** 2002. Identification of *Pantoea stewartii* subsp. *stewartii* by PCR and strain differentiation by PFGE. Plant Disease **86:**304-311.
8. **Ravin NV.** 2011. N15: the linear phage-plasmid. Plasmid **65:**102-109.
9. **Angiuoli SV, Gussman A, Klimke W, Cochrane G, Field D, Garrity G, Kodira CD, Kyrpides N, Madupu R, Markowitz V, Tatusova T, Thomson N, White O.** 2008. Toward an online repository of Standard Operating Procedures (SOPs) for (meta)genomic annotation. Omics : A Journal of Integrative Biology **12:**137-141.
10. **Tatusova T, Ciufo S, Fedorov B, O'Neill K, Tolstoy I.** 2014. RefSeq microbial genomes database: new representation and annotation strategy. Nucleic Acids Research **42:**D553-559.
11. **Correa VR, Majerczak DR, Ammar el D, Merighi M, Pratt RC, Hogenhout SA, Coplin DL, Redinbaugh MG.** 2012. The bacterium *Pantoea stewartii* uses two different type III secretion systems to colonize its plant host and insect vector. Applied and Environmental Microbiology **78:**6327-6336.

PART II: Expanded Analysis of the *Pantoea stewartii* subsp. *stewartii* DC283 Complete Genome Reveals Plasmid-borne Virulence Factors

ABSTRACT

Pantoea stewartii subsp. *stewartii*, a Gram-negative proteobacterium, native to North America, causes Stewart's wilt disease in sweet corn. Bacterial transmission to plants primarily occurs during the feeding process of the corn flea beetle insect vector. *P. stewartii* DC283 is the wild-type reference strain primarily used to study pathogenesis. Previously the complete genome of *P. stewartii* was released. Here, the method whereby the genome was assembled is described in greater detail. Data from a mate-pair library preparation with 3.5 kilobase insert size and high-throughput sequencing from the MiSeq Illumina platform, together with the available incomplete genome sequence of AHIE00000000.1 (containing 65 contigs) was used. This work resulted in the complete assembly of one circular chromosome, ten circular plasmids and one linear phage from *P. stewartii* DC283. A high number of sequences encoding repetitive transposases (> 400) were found in the complete genome. The separation of plasmids from genomic DNA revealed that two Type III secretion systems in *P. stewartii* DC283 are located on two separate mega-plasmids. Interestingly, the assembly identified a previously unknown 66-kb region in a location interior to contig 8 in the previous reference genome. Overall, a novel approach was successfully utilized to fully assemble a prokaryotic genome that contains large numbers of repetitive sequences and multiple plasmids, which resulted in some interesting biological findings.

INTRODUCTION

Pantoea stewartii subsp. *stewartii* (referred to herein as *P. stewartii*), a Gram-negative gamma-proteobacterium, native to North America, causes Stewart's wilt disease in sweet corn (2). *P. stewartii* belongs to the *Enterobacteriaceae* family containing plant-associate enterics (e.g. *Erwinia*, *Dickeya*, and *Pectobacterium* sp.) and human enteric and plant-associated pathogens (e.g. *Escherichia* and *Salmonella* sp.) (12). *P. stewartii* can colonize and grow to high cell density in the xylem of sweet corn, after being transmitted by the corn flea beetle, *Chaetocnema pulicaria* (13). The symptoms of Stewart's wilt include water-soaked lesions when the infection occurs in the apoplast space at the early stages of the disease, wilting if the infection becomes systemic through the xylem vessels, and death if the plants were infected at their seedling phase (2, 14).

P. stewartii is a good model for studying gene regulation by quorum sensing (QS, a method of cell-cell communication among eubacteria), host-pathogen interactions, and bacterial surface motility. The QS master regulator in *P. stewartii*, EsaR, represents a distinctive subfamily of LuxR-type proteins with regard to structure and function (15, 16). *P. stewartii* interacts with both plant and insect hosts making its life-style interesting to investigate. Further, it is a risk-group category one organism which is ideal for genetic manipulation with multiple genetic approaches such as reverse genetics, transcriptomic analysis and transposon mutagenesis sequencing (Tn-Seq). *P. stewartii* possesses surface motility, without any detectable swimming activity, which may contribute to the migration in the xylem and formation of biofilm. In the past, this organism was classified as non-motile, its surface motility was only recently recognized and shown to be controlled by QS and require flagella (17).

P. stewartii DC283, a nalidixic acid resistant mutant of the original *Zea mays* 1976 isolate SS104 (6, 7), is used as a wild-type reference strain to study pathogenesis of this phytopathogen. A draft genome assembly of *P. stewartii* DC283 was issued in 2012 with 65 contigs (NCBI GenBank: AHIE00000000.1). This enabled the initial genetic approaches to study this pathogen at larger scale, such as studying the whole transcriptome of the microorganism with RNA-Seq (5, 18). Since 2014, additional draft *Pantoea stewartii* genomes from different subspecies have been released but none were fully assembled (19). The difficulty of the whole genome assembly in *P. stewartii* DC283 could be explained in part by the constraints of the sequencing technologies available at the time, as well as the high number of plasmids naturally present in this bacterium (6), and the high amount of transposable elements in its genome (NCBI GenBank: AHIE00000000.1).

There have been numerous studies focused on understanding the pathogenicity of this phytopathogen in order to counteract its effects on the plants and/or to break the infection cycle to protect the crop (2, 4, 11, 17, 20-24). Two main systems in *P. stewartii* have been characterized as the underlying mechanisms of wilt-disease symptoms. First, the Hrp-type III secretion system (T3SS) is linked to the water-soaked lesions that occur during the early stages of the disease (2, 22). Indeed, *P. stewartii* utilizes two T3SS to maintain its interaction in different hosts, one for corn and the other for the corn flea beetles (11, 25). Second, stewartan, a specific extracellular polysaccharide (EPS) produced by *P. stewartii* (26), is considered to be a main pathogenesis factor responsible for a number of disease symptoms such as vascular streaking, bacterial oozing and wilting (2, 27, 28). Stewartan production is controlled by QS regulation at low cell density via the repression of RcsA, the transcription activator controlling capsule production (29). QS is a bacterial communication system that controls gene expression in

response to cell densities (30). In *P. stewartii*, the QS signal is an acyl homoserine lactone (AHL) produced by AHL synthase, EsaI, a homolog of LuxI (1). This diffusible AHL interacts with EsaR, a homolog of LuxR, at high cell densities and prevents DNA-binding by the transcription regulator. Thus, genes repressed or activated at low cell densities are derepressed or deactivated, respectively (16, 31). QS is also known to regulate the surface motility of *P. stewartii* which contributed to the virulence in the plants (17); however, the mechanism of this regulation remains unclear. LrhA, a direct target of EsaR (32), promotes surface motility and represses expression of *rcsA* (5). To better facilitate genomic-level analysis of *P. stewartii*, including its transcriptome and genes essential to growth as determined by Tn-Seq, generation of a complete genome sequence became imperative.

DNA sequencing has rapidly advanced with new technological developments over the last decade (33). The current state of the art of DNA sequencing is termed next-generation sequencing (NGS) in which the method allows parallel sequencing of large numbers of relatively small fragments of DNA (reads) and assembly into large contiguous sequences (contigs) (34). Illumina is one of the four manufacturers providing currently available platforms to perform NGS (35). Illumina applies the optical detection of clonal amplification to determine the DNA sequence during the synthesis of the prepared templates (http://www.illumina.com). Mate-pair library preparation with large insertion size is a recent tool to improve the ability of the Illumina platform (36, 37) to complete the assembly of genomes, since it permits the extension and linkage of contigs terminated by long repetitive sequences (38). For this study 250 bp reads with a 3.5 kb insert were successfully used to generate linked contigs that were assembled into the complete genome of *P. stewartii* DC283.

MATERIALS AND METHODS

Library preparation and Illumina sequencing.

The overall workflow is depicted in Figure 3.1. *P. stewartii* DC283 was grown in 5 ml Luria Bertani (LB) medium (10 g/l tryptone, 5 g/l yeast extract, and 5 g/l NaCl) overnight and the cell pellet was harvested for genome extraction using a QIAgen DNeasy Blood & Tissue Kit per the manufacturer's recommendations for Gram-negative bacteria protocol. On-column RNase treatment was applied before DNA elution to yield high-quality DNA for mate-pair library preparation and sequencing with pair-end sequencing technology (250 bp mate-pair reads with a 3500 bp insert) on an Illumina MiSeq platform at the Genomics Research lab at the Biocomplexity Institute (Virginia Tech, VA). The mate-pair library construction was performed using the Illumina Nextera Mate-Pair protocol, gel plus method (Fig 3.2). Size selection was done on a Sage Science Pippin Prep using a 0.75% gel. Then ~3500 bp DNA fragments were circularization and sheared on the Covaris m220 to select for tagged-fragments of DNA. Finally, the library was quantified using qPCR (Kapa Kit) and then pooled for sequencing on the MiSeq using a v2 500 cycle kit set to do 2X 250PE.

Bioinformatics analysis.

Raw Illumina sequencing data was generated by the core sequencing facility in the form of fastq files. These fastq files were imported into Geneious V9.1.2 (Biomatters Ltd.) for further analysis. The mate-reads were first matched together using the 'Set Paired reads' function and then assembled using the 'Map to Reference' function to the 65 contigs from AHIE00000000.1 as the reference. Unmapped mate-pair reads were then *de novo* assembled to identify missing sequences and rearrangements of the reference contigs. Finally, the end-sequences of all of the remaining contigs were extended ~3,500 bp by *de novo* assembling mate-pair reads where one

mate maps to the last 3,500 bp of the contig. The Geneious *de novo* assembly was used on unmapped reads to identify missing sequences and link the gaps. Plasmids were identified and separated from the chromosomal DNA based on their average read coverage, as well as their ability to be circularized at two ends of their overlapping sequences.

Assembly confirmation via PCR.

Traditional PCR reactions were utilized to verify some of the connections between the contigs during the assembly process. In particular, PCR reactions were designed to verify the existence of the novel 66-kilobase (kb) sequence that resulted from the *de novo* assembly. Primers were designed to have a similar melting temperature (approximately 60^{o}C) with a length between 20-30 bp in order to run them in parallel. Two pairs of forward and reverse primers were used to verify the connection between the 66-kb sequence with the interior of contig 8 (AHIE01000008) (PSG-1F/ PSG-1R & PSG-2F/ PSG-2R, Table 3.1). One*Taq*® 2X Master Mix with Standard Buffer (New England Biolabs, USA) was used with the reaction volume of 15 µl and 667 nM of each primer. The thermocycler settings were 30 s at 94^{o}C, 30 s at 55^{o}C, and 1m 45s 72^{o}C for 30 cycles. Amplicons were visualized on a 1% agarose gel and then extracted prior to sequencing to confirm the generation of the proper products.

Another set of primers (PSG-3F/ PSG-3R, Table 3.1) was used to confirm the connection between two separate contigs (i.e. contig 8 and contig 49) after bioinformatics analysis to verify the accuracy of the contig end extension using One*Taq* reactions as described above.

A QIAGEN® LongRange PCR kit (QiaGen, USA) was used to amplify the four identified rRNA operons through bioinformatics analysis, which span approximately 6 kb, and to assist in the determination of the direction of connections of the last three large assembled

chromosomal contigs whose ends contain multiple transposons, in addition one of these ends also had a rRNA operon. See Table 3.1 for the primer sequences, which were designed as described above. A reaction volume of 25 µl and 400 nM of each primer were used per the manufacturer's instruction. The thermocycler settings were 15 s at 93°C, 30 s at 55°C, and 7 m at 68°C for 35 cycles. Amplicons were again visualized on a 1% agarose gel and then extracted prior to sequencing.

Nucleotide sequence accession number.

As previously published in *Genome Announcements* (39), the complete genome was annotated using the NCBI prokaryotic annotation pipeline and the complete genome assembly of *P. stewartii* DC283 is available in GenBank under accession no. CP017581-CP017592.

Development of PCR method to screen for the presence of the 66-kb region.

A multiplex PCR reaction with three primers for rapid detection of the missing 66-kb region was developed for the analysis of new deletion strain constructs. This reaction contained primers 667 nM PSG-1F, 667 nM PSG-2F and 667 nM PSG-2R in a 15 µl final volume using One*Taq*® 2X Master Mix with Standard Buffer (NEB). The thermocycler settings were 30 s at 94°C, 30 s at 55°C, and 1 m 45 s at 72°C for 30 cycles. The amplicon was visualized on a 1% agarose gel to determine its size. This corresponds to the presence (1736 bp, product of PSG-2F and PSG-2R) or absence (1324 bp, product of PSG-1F and PSG-2R) of the 66-kb region. The reaction was then repeated with PSG-1F and PSG-1R to confirm the presence or absence of the 66-kb sequence.

RESULTS

Comparison of the previous incomplete and new complete genome sequences of *P. stewartii* DC283.

The complete genome of *P. stewartii* DC283 consists of 5,314,092 bp (53.8% G+C content) with 5,625 coding sequences, 21 rRNAs and 73 tRNAs (39). This genome includes one circular chromosome, ten circular plasmids and one linear phage (Table 3.2). The previous genome assembly (AHIE00000000.1) consisting of 65 contigs was complicated by the large number of repetitive transposon sequences that prematurely terminated contigs and introduced ambiguities in the assembly. The complete assembly identified 444 repetitive transposon sequences often spanning ~1,500 bp that could nevertheless be bridged using the mate-pair reads with 3,500 bp inserts. These mate-pair reads linked the ends of the previous 65 contigs and led to rearrangements of the sequences in some of them. Table 3.3 summarizes some major features of the sequencing results for the complete *P. stewartii* DC283 genome generated in this study in comparison with the initial incomplete reference sequence (AHIE00000000.1).

Two type III secretion systems in *P. stewartii* are located on two separate mega-plasmids.

To facilitate the building of connections in the main chromosomal DNA during the process of contig end extension, using multiple rounds of mate-pair recognition and mapping to the remaining contigs of the incomplete genome, sequences that could be circularized were separated to access their function as plasmids. These sequences were then used to blast with the database on NCBI to confirm their identities. Several of these sequences were highly similar to known *P. stewartii* plasmids sequences, e.g. contig 52 (AHIE01000052) was found to be similar to pSW100 (40) and is now known as pDSJ01 (4,277 bp), contig 62 (AHIE01000062) was found to be similar to pSW200 (41) and is now known as pDSJ02 (4,368 bp), and contig 65

(AHIE01000065) was found to be similar to pSW800 (42) and is now known as pDSJ05 (34,447 bp). Preliminary annotation was performed by rapid annotation using subsystem technology (RAST, http://rast.nmpdr.org/). This approach was used to identify genes involving in the replication of plasmids (eg. *repA*) in the circular sequences prior to submission to NCBI for final annotation and publication. In addition, the amount of read coverage determined during assembly was used to calculate the average copy number of each plasmid in *P. stewartii* DC283. Ten separate plasmids with their copy numbers were identified and renamed in order from smallest to largest according to their molecular mass (Table 3.2). Plasmids pDSJ01 and pDSJ02 exist as mediate-level copy numbers in *P. stewartii* while the other plasmids and the linear phage-plasmid present as low copies. The three mega-plasmids (size above 100 kb) have the lowest copy numbers, from 1-3.

The separation of the plasmids from the genomic DNA and their annotation revealed that the two T3SS in *P. stewartii* DC283 are located on two separate mega-plasmids. Specifically, genes related to the T3SS needed for the invasion of the insect host and colonization of the plant host (11) are located in plasmid pDSJ08 and plasmid pDSJ10, respectively, in the complete genome.

Identification of a phage N15-like linear phage plasmid of *P. stewartii.*

There are a high number of sequence reads mapped to contig 47 (AHIE01000047, with 46,532 bp) in the draft genome having a coverage of ~1000X, which may correspond to approximately six copies of this sequence in the *P. stewartii* DC283 genome. This contig also contains genes encoding a partitioning system usually found in plasmid sequences. However, the contig end extension using mate-pair reads did not result in a connection with any other contigs or self-circularization. This suggests that the sequence is a linear extrachromosomal DNA

element present in multiple copies. The annotation of this contig also showed a high number of phage-related coding sequences. In addition, data from the blastx function on the NCBI website against non-redundant protein sequences (nr) database for this sequence showed some level of similarity to coding sequences in a N15 prophage in *Escherichia coli* (8). Therefore, this contig-end fixed sequence has been renamed as ppDSJ01 (47,186 bp) to represent a linear phage plasmid of *P. stewartii* DC283 (Table 3.2).

Multiple chromosomal prophage sequences found in *P. stewartii* genome.

When the sequenced reads were mapped to draft contig 8 (AHIE01000008), an interesting mapping coverage was observed (Fig 3.4) from position ~140,000 to ~200,000. The average coverage of the entire contig is ~320X with the majority of the sequence having coverage at ~200X which is the baseline of coverage of the chromosomal DNA as a single copy in a cell. However, the specific region of DNA shown in Figure 3.4 has three regions of elevated coverage. Two of these regions have around two times the coverage of the chromosomal DNA with a region of up to ~16 times the chromosome coverage in the middle region spanning between positions of ~165,000 and ~180,000 of contig 8. The entire elevated region spans from position ~2,032,000 to ~2,089,000 of the complete *P. stewartii* chromosome (CP017581) with the middle region spanning between positions of ~2,053,000 and ~2,066,000. The annotation of these elevated coverage regions showed that they contain genes related to bacteriophage sequences. It was not possible to distinguish whether these sequences belong to tandem sequence duplications in the chromosome in this region or whether some of them are also present in the cell as extrachromosomal linear phage DNAs.

Similarly, another region spanning from ~2,411,000 to ~2,463,000 of the complete *P. stewartii* chromosome DNA molecule was recognized as duplication in the sequence coverage

suggesting either tandem repeat of the ~51,000 bp region or the extrachromosomal existence of a linear DNA molecule with the same sequence chromosomally integrated. The annotation of genes included inside this region showed a large amount of hypothetical proteins and phage-related proteins.

Seven contigs in the reference genome are absent in the complete *P. stewartii* genome.

Surprisingly, there are seven small contigs in the incomplete genome with no sequence reads mapped to their sequences on the new complete genome. The total length of these contigs is 23,196 bp with 17 genes (Table 3.4). This content of DNA makes up only 0.4 % of the total DNA length of the new assembly of the *P. stewartii* DC283 genome. It could not be determined whether these contigs are real and reflect a discrepancy between the two DNA genome sequences of *P. stewartii* DC283, or they are results of sequencing and assembly errors. However, the annotation of the genes located on these contigs suggested that they may be part of extrachromosomal DNA such as plasmids or mobile elements like transposons (Table 3.4) in the previous draft genome.

Identification of a novel region of *P. stewartii* genome.

Interestingly, the assembly of the new sequencing data resulted in the identification of a novel region, 66 kb in length containing 68 genes, in the interior of contig 8 of the previous reference genome. PCR reactions were used to confirm the existence of this region inside this contig. These PCR results followed by Sanger sequencing of the PCR products have confirmed the integrity of the bioinformatics analysis using *de novo* assembly of the unmapped reads and the contig end extension methods. This region is now located from 1,862,325 to 1,928,315 of the *P. stewartii* main chromosome (CP017581).

Additionally, the connection between the two ends of contig 8 (containing the 66-kb region) and contig 49 into a larger contig using bioinformatic extension was verified using PCR and sequencing.

Analysis of chromosome stability among genetic mutants of *P. stewartii* DC283.

Two strains of wild-type *P. stewartii* DC283 (one in our lab and one from Dr. David Mackey's lab at the Ohio State University) and all mutants generated in our lab were tested for the loss of the novel 66-kb region of the genome. A multiplex PCR reaction with primers designed to confirm the connection of this region to the interior of contig 8 was used to screen for the presence or absence of this 66-kb region. The results of the screening PCR showed that only the published *rcsA* strain and its complementation strain (5) constructed in our lab did not have the newly identified 66-kb region whereas it is present in the rest of tested strains.

DISCUSSION

A novel approach involving mate-pair library preparation with pair-end Illumina sequencing technology was successfully utilized to fully assemble the *P. stewartii* prokaryotic genome that contains large numbers of repetitive sequences and multiple plasmids. This whole assembly of the *P. stewartii* genome has revealed a rich content of extrachromosomal DNA. Two small plasmids, each ~4.5 kb, were present at a medium copy number whereas the other plasmids are maintained in less than 10 copies per cell. The larger the size of a plasmid, the less abundant it seems to be in *P. stewartii* cells. This negative correlation between plasmid size and their copy number in a bacterial cell was previously experimentally determined in *Bacillus thuringiensis* YBT-1520 (43). The genome of *Bacillus thuringiensis* YBT-1520 contains 11 plasmids (between 2 kb and 416 kb in size) with their copy numbers negatively correlated to their molecular mass.

This general trend may correspond to the energetic burden of carrying a plasmid and the maintaining of its copy number inside the cells (44, 45).

The two T3SS used by *P. stewartii* to independently colonize two different hosts and maintain its life cycle between the plants and the insect vectors (11) are now known to be maintained on two mega-plasmids (pDSJ010 and pDSJ008), respectively. The largest *P. stewartii* plasmid, pDSJ010, contains the plant colonizing T3SS and it is the universal plasmid LPP-1 in the genus *Pantoea* (46). LPP-1 plasmids, derived from an ancestral plasmid, contain a large repertoire of proteins that contribute to the adaptation of species of the genus to their various niches and their specialization as beneficial biocontrol agents, harmless saprophytes, or pathogens (46). Other plasmids carried by *P. stewartii* may also play important roles in many cellular biophysical processes as well as colonization and pathogenesis within multiple hosts. The complete sequence for these plasmids will facilitate the utilization of bioinformatic approaches to understand their roles in relation to host bacterium.

This project also resulted in the identification of the extrachromosomal presence of a linear plasmid phage in multiple copies which is similar to the lambdoid phage N15 of *Escherichia coli* (8). Additional phage elements found in the *P. stewartii* genome at positions ~2,032,000 - ~2,089,000 and ~2,411,000 - ~2,463,000 could be intrachromosomally incorporated in tandem and/or present as linear plasmid phages with several copies per cell. Further investigation is needed to clarify this issue in the assembly as well as to understand the role of phage elements inside this bacterial host.

Interestingly, a small amount of DNA sequence (~23 kb) from the incomplete genome was not present in the newly assembled genome of *P. stewartii* DC283. This might have arisen through technical discrepancies between the methods used to generate the two genome

sequences. Alternatively, these missing sequences may belong to unstable DNA molecules that led to genetic content differences among the reference strains of *P. stewartii* DC283 after they were domesticated in multiple laboratories for decades. The second reason seems to be probable for at least some of these sequences when the annotation was taken into account. The majority of these genes code for conjugative components of plasmid transfers, phage-related proteins, and transposon's elements (Table 3.4).

The recognition of the 66-kb region missing in the reference genome and a mutation strain (*rscA* deletion strain (5)) generated from our laboratory led to two findings. The first one is the existence of this region in the genome which may be involved in several different metabolic pathways since it contains 68 coding sequences. This also raised a concern about the stability of the genetic material within this species. A rapid PCR detection method was developed in this study to screen for the presence of this region among genetically modified strains in our laboratory and strains from a collaborator (47). This preliminary survey resulted in the confirmation of the presence of this region in all test strains with the exception of the *rcsA* deletion strain. This provides confidence in the overall stability of this region among *P. stewartii* laboratory strains. However, the frequency of elimination of this region, what triggers it and the consequential biological effects of this region on the bacteria during infection remain unknown.

In conclusion, this project has provided a complete genome of *P. stewartii*, a phytopathogen with a complex genome containing multiple plasmids, phages and other mobile elements such as transposons. This insight undoubtedly will facilitate further understanding of this pathogen as well as the *Pantoea* genus using bioinformatic analysis and high-throughput sequencing technology.

ACKNOWLEDGMENTS

We thank Susanne von Bodman for sharing her stock strain of *P. stewartii* DC283 with us. We thank David Mackey for sharing his stock strain of *P. stewartii* DC283 plus a *wtsE* mutant and its complementation strain with us. This work was funded by the Fralin Life Science Institute (RJV) and Biological Sciences Department (AMS), Virginia Tech. Mate-pair library construction and Illumina sequencing were conducted by Genomics Research Laboratory at the Biocomplexity Institute at Virginia Tech.

REFERENCES

1. **von Bodman SB, Bauer WD, Coplin DL.** 2003. Quorum sensing in plant-pathogenic bacteria. Annual Review of Phytopathology **41:**455-482.
2. **Roper MC.** 2011. *Pantoea stewartii* subsp. *stewartii*: lessons learned from a xylem-dwelling pathogen of sweet corn. Molecular Plant Pathology **12:**628-637.
3. **Ham JH, Majerczak D, Ewert S, Sreerekha MV, Mackey D, Coplin D.** 2008. WtsE, an AvrE-family type III effector protein of *Pantoea stewartii* subsp. *stewartii*, causes cell death in non-host plants. Molecular Plant Pathology **9:**633-643.
4. **Burbank L, Roper MC.** 2014. OxyR and SoxR modulate the inducible oxidative stress response and are implicated during different stages of infection for the bacterial phytopathogen *Pantoea stewartii* subsp. *stewartii*. Molecular Plant-Microbe Interactions **27:**479-490.
5. **Kernell Burke A, Duong DA, Jensen RV, Stevens AM.** 2015. Analyzing the transcriptomes of two quorum-sensing controlled transcription factors, RcsA and LrhA, important for *Pantoea stewartii* virulence. PLOS ONE **10:**e0145358.
6. **Coplin DL, Rowan RG, Chisholm DA, Whitmoyer RE.** 1981. Characterization of plasmids in *Erwinia stewartii*. Applied and Environmental Microbiology **42:**599-604.
7. **Coplin DL, Majerczak DR, Zhang Y, Kim W, Jock S, Geider K.** 2002. Identification of *Pantoea stewartii* subsp. *stewartii* by PCR and strain differentiation by PFGE. Plant Disease **86:**304-311.
8. **Ravin NV.** 2011. N15: the linear phage-plasmid. Plasmid **65:**102-109.
9. **Angiuoli SV, Gussman A, Klimke W, Cochrane G, Field D, Garrity G, Kodira CD, Kyrpides N, Madupu R, Markowitz V, Tatusova T, Thomson N, White O.** 2008. Toward an online repository of Standard Operating Procedures (SOPs) for (meta)genomic annotation. Omics : A Journal of Integrative Biology **12:**137-141.
10. **Tatusova T, Ciufo S, Fedorov B, O'Neill K, Tolstoy I.** 2014. RefSeq microbial genomes database: new representation and annotation strategy. Nucleic Acids Research **42:**D553-559.
11. **Correa VR, Majerczak DR, Ammar el D, Merighi M, Pratt RC, Hogenhout SA, Coplin DL, Redinbaugh MG.** 2012. The bacterium *Pantoea stewartii* uses two different type III secretion systems to colonize its plant host and insect vector. Applied and Environmental Microbiology **78:**6327-6336.
12. **Walterson AM, Stavrinides J.** 2015. *Pantoea:* insights into a highly versatile and diverse genus within the *Enterobacteriaceae*. FEMS Microbiology Reviews **39:**968-984.
13. **Esker PD, Nutter FW.** 2002. Assessing the risk of Stewart's disease of corn through improved knowledge of the role of the corn flea beetle vector. Phytopathology **92:**668-670.
14. **Braun EJ.** 1982. Ultrastructural investigation of resistant and susceptible maize inbreds infected with *Erwinia stewartii*. Phytopathology **72:**159-166.
15. **Minogue TD, Wehland-von Trebra M, Bernhard F, von Bodman SB.** 2002. The autoregulatory role of EsaR, a quorum-sensing regulator in *Pantoea stewartii* ssp. *stewartii*: evidence for a repressor function. Molecular Microbiology **44:**1625-1635.
16. **Schu DJ, Carlier AL, Jamison KP, von Bodman S, Stevens AM.** 2009. Structure/function analysis of the *Pantoea stewartii* quorum-sensing regulator EsaR as an activator of transcription. Journal of Bacteriology **191:**7402-7409.
17. **Herrera CM, Koutsoudis MD, Wang X, von Bodman SB.** 2008. *Pantoea stewartii* subsp. *stewartii* exhibits surface motility, which is a critical aspect of Stewart's wilt disease development on maize. Molecular Plant-Microbe Interactions **21:**1359-1370.

18. **Ramachandran R, Burke AK, Cormier G, Jensen RV, Stevens AM.** 2014. Transcriptome-based analysis of the *Pantoea stewartii* quorum-sensing regulon and identification of EsaR direct targets. Applied and Environmental Microbiology **80:**5790-5800.
19. **De Maayer P, Aliyu H, Vikram S, Blom J, Duffy B, Cowan DA, Smits THM, Venter SN, Coutinho TA.** 2017. Phylogenomic, pan-genomic, pathogenomic and evolutionary genomic insights into the agronomically relevant Enterobacteria *Pantoea ananatis* and *Pantoea stewartii*. Frontiers in Microbiology **8:**1755.
20. **von Bodman SB, Majerczak DR, Coplin DL.** 1998. A negative regulator mediates quorum-sensing control of exopolysaccharide production in *Pantoea stewartii* subsp. *stewartii*. Proceedings of the National Academy of Sciences of the United States of America **95:**7687-7692.
21. **Minogue TD, Carlier AL, Koutsoudis MD, von Bodman SB.** 2005. The cell density-dependent expression of stewartan exopolysaccharide in *Pantoea stewartii* ssp. *stewartii* is a function of EsaR-mediated repression of the *rcsA* gene. Molecular Microbiology **56:**189-203.
22. **Ham JH, Majerczak DR, Arroyo-Rodriguez AS, Mackey DM, Coplin DL.** 2006. WtsE, an AvrE-family effector protein from *Pantoea stewartii* subsp. *stewartii*, causes disease-associated cell death in corn and requires a chaperone protein for stability. Molecular Plant-Microbe Interactions **19:**1092-1102.
23. **Menelas B, Block CC, Esker PD, Nutter FW, Jr.** 2006. Quantifying the feeding periods required by corn flea beetles to acquire and transmit *Pantoea stewartii*. Plant Disease **90:**319-324.
24. **Burbank L, Mohammadi M, Roper MC.** 2014. Siderophore-mediated iron acquisition influences motility and is required for full virulence for the xylem-dwelling bacterial phytopathogen, *Pantoea stewartii* subsp. *stewartii*. Applied and Environmental Microbiology.
25. **Packard H, Kernell Burke A, Jensen RV, Stevens AM.** 2017. Analysis of the *in planta* transcriptome expressed by the corn pathogen *Pantoea stewartii* subsp. *stewartii* via RNA-Seq. PeerJ **5:**e3237.
26. **Nimtz M, Mort A, Wray V, Domke T, Zhang Y, Coplin DL, Geider K.** 1996. Structure of stewartan, the capsular exopolysaccharide from the corn pathogen *Erwinia stewartii*. Carbohydrate Research **288:**189-201.
27. **Beck von Bodman S, Farrand SK.** 1995. Capsular polysaccharide biosynthesis and pathogenicity in *Erwinia stewartii* require induction by an N-acylhomoserine lactone autoinducer. Journal of Bacteriology **177:**5000-5008.
28. **Dolph PJ, Majerczak DR, Coplin DL.** 1988. Characterization of a gene cluster for exopolysaccharide biosynthesis and virulence in *Erwinia stewartii*. Journal of Bacteriology **170:**865-871.
29. **Carlier A, Burbank L, von Bodman SB.** 2009. Identification and characterization of three novel EsaI/EsaR quorum-sensing controlled stewartan exopolysaccharide biosynthetic genes in *Pantoea stewartii* ssp. *stewartii*. Molecular Microbiology **74:**903-913.
30. **de Kievit TR, Iglewski BH.** 2000. Bacterial quorum sensing in pathogenic relationships. Infection and Immunity **68:**4839-4849.
31. **von Bodman SB, Ball JK, Faini MA, Herrera CM, Minogue TD, Urbanowski ML, Stevens AM.** 2003. The quorum sensing negative regulators EsaR and ExpR(Ecc), homologues within the LuxR family, retain the ability to function as activators of transcription. Journal of Bacteriology **185:**7001-7007.
32. **Ramachandran R, Stevens AM.** 2013. Proteomic analysis of the quorum-sensing regulon in *Pantoea stewartii* and identification of direct targets of EsaR. Applied and Environmental Microbiology **79:**6244-6252.
33. **Pareek CS, Smoczynski R, Tretyn A.** 2011. Sequencing technologies and genome sequencing. Journal of Applied Genetics **52:**413-435.

34. **Rizzo JM, Buck MJ.** 2012. Key principles and clinical applications of "next-generation" DNA sequencing. Cancer Prevention Research **5:**887-900.
35. **Levy SE, Myers RM.** 2016. Advancements in next-generation sequencing. Annual Review of Genomics and Human Genetics **17:**95-115.
36. **Srivastava A, Philip VM, Greenstein I, Rowe LB, Barter M, Lutz C, Reinholdt LG.** 2014. Discovery of transgene insertion sites by high throughput sequencing of mate pair libraries. BMC Genomics **15:**367.
37. **Jiao X, Hooper SD, Djureinovic T, Larsson C, Warnberg F, Tellgren-Roth C, Botling J, Sjoblom T.** 2013. Gene rearrangements in hormone receptor negative breast cancers revealed by mate pair sequencing. BMC Genomics **14:**165.
38. **Mardis E, McCombie WR.** 2017. Preparation of a 3-kb mate-pair library for Illumina sequencing. Cold Spring Harbor Protocols **2017:**pdb prot094656.
39. **Duong DA, Stevens AM, Jensen RV.** 2017. Complete genome assembly of *Pantoea stewartii* subsp. *stewartii* DC283, a corn pathogen. Genome Announcements **5**: e00435-17.
40. **Fu JF, Chang HC, Chen YM, Chang YS, Liu ST.** 1995. Sequence analysis of an *Erwinia stewartii* plasmid, pSW100. Plasmid **34:**75-84.
41. **Fu JF, Hu JM, Chang YS, Liu ST.** 1998. Isolation and characterization of plasmid pSW200 from *Erwinia stewartii*. Plasmid **40:**100-112.
42. **Wu CY, Fu JF, Liu ST.** 2001. The replicon of pSW800 from *Pantoea stewartii*. Microbiology **147:**2757-2767.
43. **Zhong C, Peng D, Ye W, Chai L, Qi J, Yu Z, Ruan L, Sun M.** 2011. Determination of plasmid copy number reveals the total plasmid DNA amount is greater than the chromosomal DNA amount in *Bacillus thuringiensis* YBT-1520. PLOS ONE **6:**e16025.
44. **del Solar G, Espinosa M.** 2000. Plasmid copy number control: an ever-growing story. Molecular Microbiology **37:**492-500.
45. **Camps M.** 2010. Modulation of ColE1-like plasmid replication for recombinant gene expression. Recent Patents on DNA & Gene Sequences **4:**58-73.
46. **De Maayer P, Chan WY, Blom J, Venter SN, Duffy B, Smits TH, Coutinho TA.** 2012. The large universal *Pantoea* plasmid LPP-1 plays a major role in biological and ecological diversification. BMC Genomics **13:**625.
47. **Asselin JE, Lin J, Perez-Quintero AL, Gentzel I, Majerczak D, Opiyo SO, Zhao W, Paek SM, Kim MG, Coplin DL, Blakeslee JJ, Mackey D.** 2015. Perturbation of maize phenylpropanoid metabolism by an AvrE family type III effector from *Pantoea stewartii*. Plant Physiology **167:**1117-1135.

Table 3.1. Primers used in this study

Name	Sequence (5' → 3')	Tm (°C)
PSG-1F	CTTAGCCAGACGACGTTTGCCG	60.6
PSG-1R	GCACATGTCGCTGATGCTGC	59.8
PSG-2F	GAGCGGGAGTAGAATCAGTAACAGTCC	59.9
PSG-2R	CGTTCCCGCTTCAGGCAACAC	61.4
PSG-3F	AGCTTTTGCCGAACTGTATACGTTGC	59.8
PSG-3R	TGTTCACCGTGGTGATATGAGACTGC	60.5
P1-32Left-F	CATCCAGTGACTGCATAAAGAAGGCC	59.7
P1-35Left-R	CGGCTTTTAAGGAGGTATACGAGCTGG	60.5
P2-36Right-F	GGCGTAAAGGACGAGGTTGTTGC	60.7
P2-30Right-R	TTGCTGAAGTTTCCAGTCGTGCC	60.1
P3_c25Right	GATGTTTGTAAGGGCGTAACGCTACAG	59.5
P3_c33Left	ATGACATGGCTGGAAAAGGCCAG	60.1
P4-1Right-F	TGACTGGGTAATAGAGAGTCTTGCGG	59.5
P4-26Right-R	GCGCGGGATTTAGATCATTACAGCC	60
F1-31Left	AGCGTTATGGCTTCCCGTTGC	60.7
F2-34Right	GCACAGTTACCAGGCAAATCGTACTG	59.8
F3-27Left	GGATGTCAAAGACGACGTCATTACGG	59.3
4R-1Left	GCATATACAGGCTCAGTTACAGCCG	59.3
5R-29Right	GCAAATCCGAAGTCAGAACGTACCC	59.7
6R-37Left	CGACGACATTCCGTTCTGACATTTACC	59.4
7R-Unique-Left37	CTCAGGTGATGTGCAGTAATGCCC	59.9
8F-Unique-Right37	CAGTTGCGCTAAGACGATCAGCG	60.3

Table 3.2. DNA components in the complete assembly of the *P. stewartii* DC283 genome.

Name	Length (bp)	Topology	Mean Coverage	Copy Number
Chromosome	4,528,215	circular	181.28	1
pDSJ01	4,277	circular	3355.36	21
pDSJ02	4,368	circular	2214.90	14
pDSJ03	13,414	circular	1414.80	9
pDSJ04	25,681	circular	952.56	6
pDSJ05	34,447	circular	593.62	4
pDSJ06	47,481	circular	414.55	3
pDSJ07	65,483	circular	408.61	2
pDSJ08	105,961	circular	440.82	3
pDSJ09	132,938	circular	277.89	2
pDSJ010	304,641	circular	163.45	1
ppDSJ01	47,186	linear	963.65	6

Table 3.3. Comparison between the incomplete (AHIE00000000.1) and the complete (CP017581-CP017592) assembly of *P. stewartii* DC283 genome

Feature	**AHIE00000000.1**	**CP017581-CP017592**
Total length (bp)	5,233,214	5,314,092
G+C %	53.8	53.8
Total DNA content	65 contigs	1 chromosome, 10 plasmids, 1 linear phage plasmid
# Proteins	4,903	4,942
# Pseudogenes	164	411
# tRNA genes	70	73
# rRNA genes	20	21
# genes annotated as encoding a transposase	361	444

Table 3.4. Seven contigs present in the incomplete genome but absent in the complete genome of *P. stewartii* DC283

Contig	Length (bp)	# Proteins	Annotations
AHIE01000051	4,392	4	conjugative transfer protein, TrbG family conjugative transfer protein, hypothetical protein, TrbE family protein
AHIE01000053	835	0	
AHIE01000054	688	0	
AHIE01000055	983	1	chloramphenicol resistance protein
AHIE01000056	1,826	2	TniQ family protein, NTP-binding protein
AHIE01000060	9,954	6	hypothetical protein, DNA replication primase, hypothetical protein, DNA topoisomerase III, TraF family plasmid transfer protein, TraG family protein
AHIE01000063	4,518	4	IS4321L family transposase, LysR family transcriptional regulator, cupin 2 conserved barrel domain protein, IS4321L family transposase
Total	**23,196**	**17**	

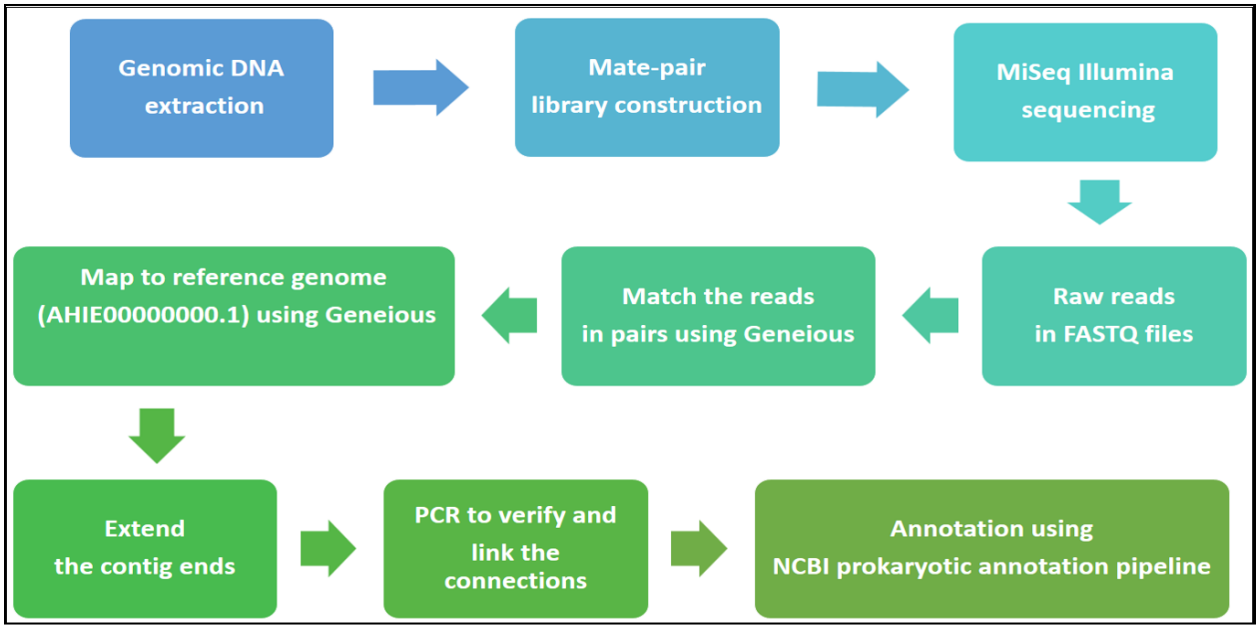

Fig 3.1. Workflow of the complete assembly of the *P. stewartii* DC283 genome. The genome was assembled using data from a mate-pair library preparation with 3.5 kb insert size and high-throughput sequencing from the MiSeq Illumina platform together with the available sequence of AHIE00000000.1.

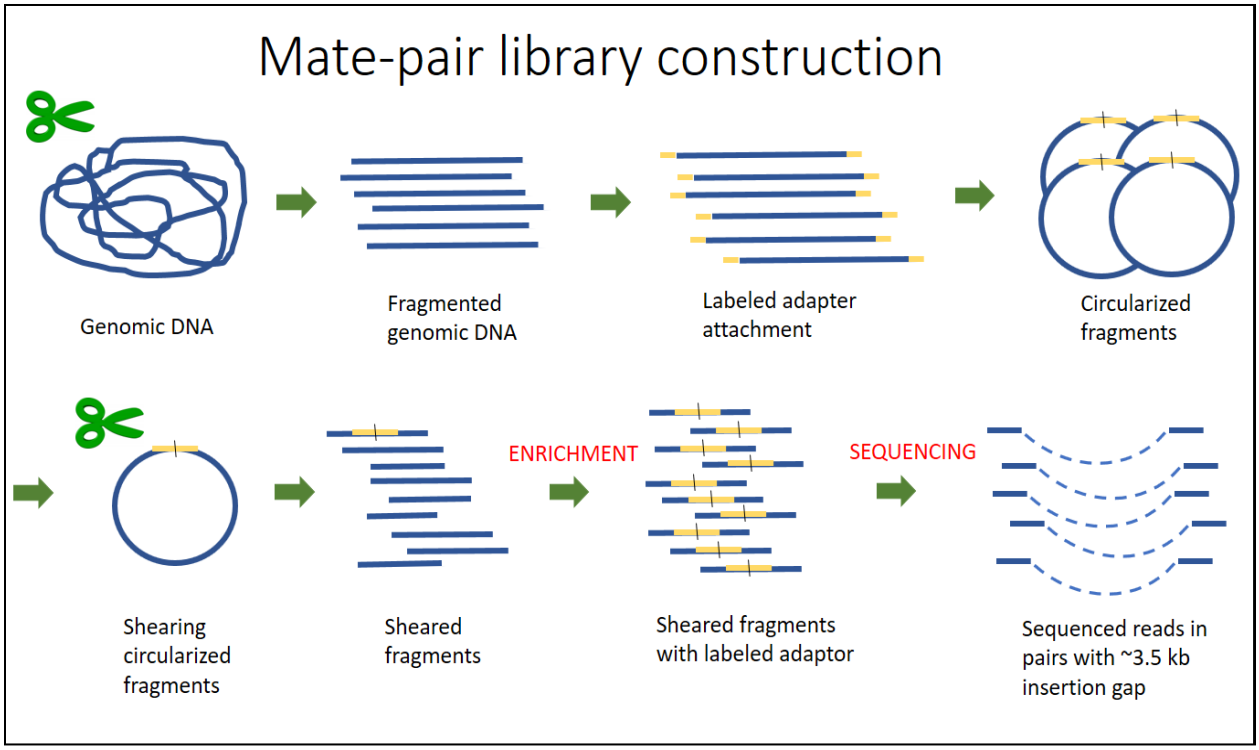

Fig 3.2. Cartoon demonstration of mate-pair library construction to generate the matched reads for the assembly process. First, genomic DNA was fragmented and selected for a desired size. These fragments were attached to labeled adapters at the two ends before circularizing. Shearing into smaller DNA fragments occurred prior to the enrichment process of those fragments containing the DNA from the two ends of the original fragments. The results of sequencing data were ~250 bp reads that can be traced back to their mates that were physically separated on the genomic DNA by ~3.5 kb.

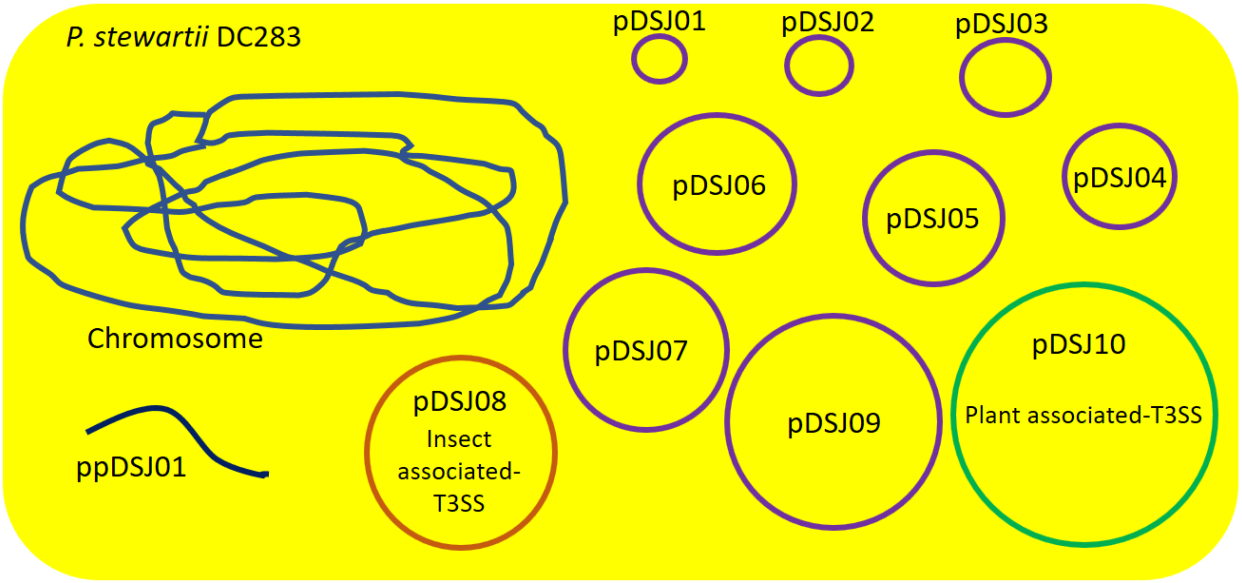

Fig 3.3. Diagram of the complete assembly of the *P. stewartii* DC283 genome. Figure is not drawn to scale.

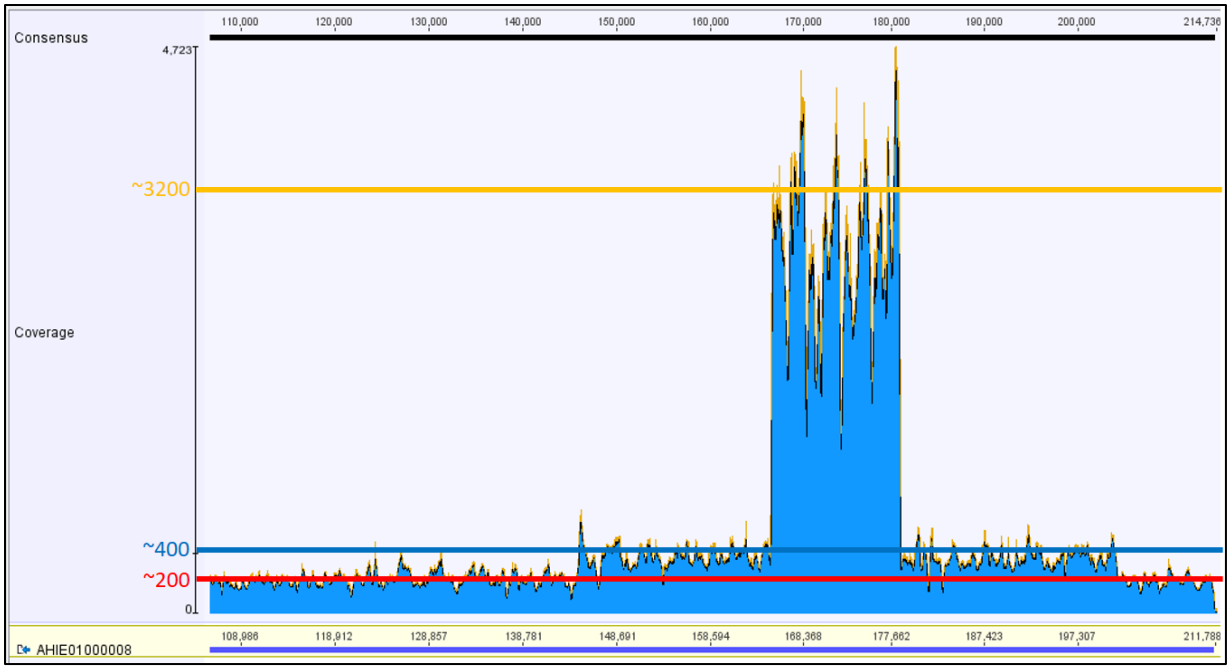

Fig 3.4. Coverage of mapped reads to a region on contig 8 (AHIE01000008) of the reference incomplete genome. The average coverage of chromosomal DNA is ~200X (the red line) while two regions, spanning ~140,000 - ~200,000, show an average of ~400X (the blue line). The yellow line indicates another region, spanning ~165,000 - ~180,000, with a coverage at ~3200X.

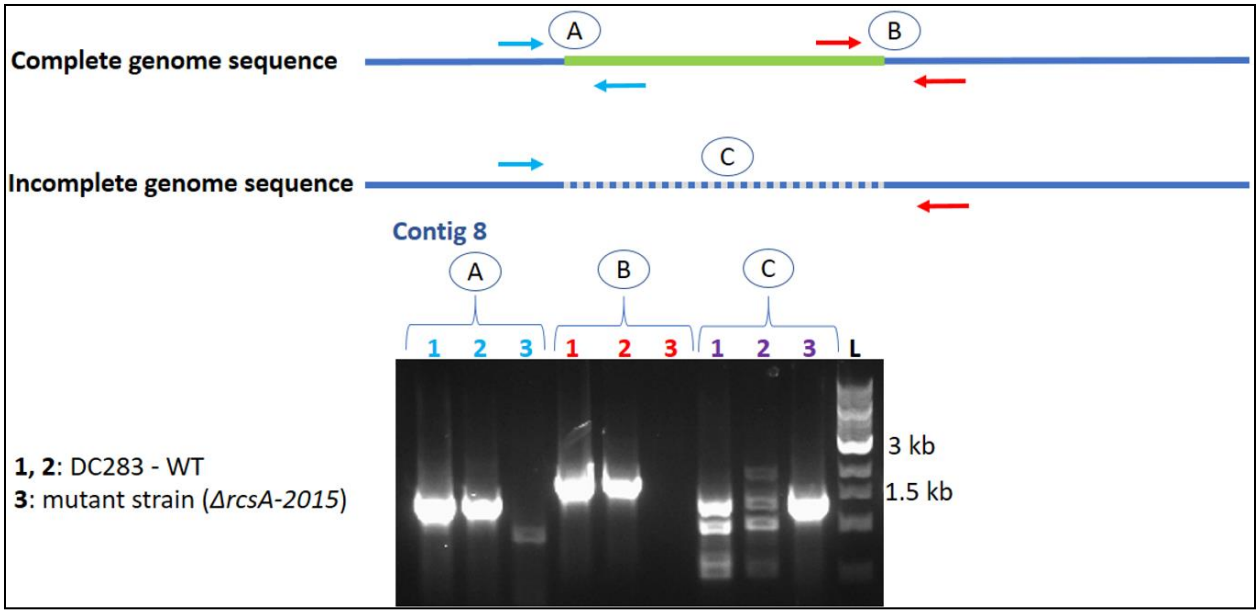

Fig 3.5. Confirmation of the identification of the *de novo* assembled 66-kb region using traditional PCR reactions. Each PCR reaction contains two primers to detect the presence of the physical connection in the tested strains. Three samples, two wild-type (WT) and a mutant (*ΔrcsA-2015*), were presented in the gel image to demonstrate the presence/absence of the appropriate specific PCR product in each sample. A: connection between the 5' end of the 66-kb region (green line) to the 3' end of the interior of contig 8 (AHIE01000008); B: connection between the 3' end of the 66-kb region (green line) to the 5' end of the interior of contig 8; C: a region interior of contig 8 in the incomplete genome sequence (AHIE00000000.1).

Chapter Four

Integrated Downstream Regulation by the Quorum-sensing Controlled Transcription Factors LrhA and RcsA Impacts Phenotypic Outputs Associated with Virulence in the Phytopathogen *Pantoea stewartii* subsp. *stewartii*

Duy An Duong and Ann M. Stevens[#]. 2017. Integrated downstream regulation by the quorum-sensing controlled transcription factors LrhA and RcsA impacts phenotypic outputs associated with virulence in the phytopathogen *Pantoea stewartii* subsp. *stewartii*. PeerJ 5:e4145. https://doi.org/10.7717/peerj.4145

#Address correspondence to Ann M. Stevens, ams@vt.edu

Running Head: LrhA and RcsA regulation in *P. stewartii*
Key words: RcsA; LrhA; *Pantoea stewartii* subsp. *stewartii*; phytopathogen; quorum sensing; transcription factor

ABSTRACT

Pantoea stewartii subsp. *stewartii* is a Gram-negative proteobacterium that causes leaf blight and Stewart's wilt disease in corn. Quorum sensing (QS) controls bacterial exopolysaccharide production that blocks water transport in the plant xylem at high bacterial densities during the later stage of the infection, resulting in wilt. At low cell density the key master QS regulator in *P. stewartii*, EsaR, directly represses *rcsA*, encoding an activator of capsule biosynthesis genes, but activates *lrhA*, encoding a transcription factor that regulates surface motility. Both RcsA and LrhA have been shown to play a role in plant virulence. In this study, additional information about the downstream targets of LrhA and its interaction with RcsA was determined. A transcriptional fusion assay revealed autorepression of LrhA in *P. stewartii* and electrophoretic mobility shift assays (EMSA) using purified LrhA confirmed that LrhA binds to its own promoter. In addition, LrhA binds to the promoter for the RcsA gene, as well as those for putative fimbrial subunits and biosurfactant production enzymes in *P. stewartii*, but not to the *flhDC* promoter, which is the main direct target of LrhA in *Escherichia coli.* This work led to a reexamination of the physiological function of RcsA in *P. stewartii* and the discovery that it also plays a role in surface motility. These findings are broadening our understanding of the coordinated regulatory cascades utilized in the phytopathogen *P. stewartii*.

INTRODUCTION

Pantoea stewartii subsp. *stewartii*, a Gram-negative rod-shaped, gamma-proteobacterium, is the causal agent of leaf blight and Stewart's wilt in susceptible varieties of *Zea mays*. It is primarily transmitted to the plant by the corn flea beetle, *Chaetocnema pulicaria* (1). After being deposited through excrement into wounds generated during insect feeding, the pathogen gains access to the leaf apoplast and causes water-soaked lesions through the Hrp-type III secretion system (2). In a second phase of the disease, the bacteria then also migrate to the xylem, where they grow to high cell density and form a biofilm that blocks water flow within the plant. This results in wilt disease and even death, if the plants were infected at the seedling phase (3). Quorum sensing (QS), a mechanism of bacterial cell density-dependent communication, controls the virulence, capsule production and surface motility of this pathogen (4, 5).

During QS, *P. stewartii* produces N-acyl homoserine lactone (AHL) signals due to the activity of EsaI, a LuxI-type protein (6). The AHL signal then interacts with the master QS regulatory protein EsaR, a LuxR homologue, when the cell density reaches a critical threshold. EsaR is a dual-level transcriptional regulator that binds to DNA at its recognition sites to either activate or repress its downstream targets at low cell density (6-8). When EsaR and AHL interact at high cell density, the EsaR-AHL complex is unable to bind to the DNA resulting in transcriptional deactivation or derepression of its target genes (9, 10). Multiple approaches have been used to identify several direct targets of EsaR, including classic genetic (11), proteome-level (12) and transcriptome-level (13) analysis. Two of these direct targets, *rcsA* and *lrhA*, are involved in plant virulence and control capsule production and surface motility, respectively (14).

EsaR directly represses the *P. stewartii rcsA* gene at low cell density, insuring precise control over the timing of capsule synthesis (8, 11, 15). At high cell density, gene activation by RcsA leads to production of stewartan, a polymer of galactose, glucose and glucuronic acid in a 3:3:1 ratio, which is the main component of the exopolysaccharide (EPS) (16). Stewartan is a primary virulence factor of *P. stewartii* (4, 11, 17). Previous work has shown that the *lrhA* gene is directly activated by EsaR at low cell density and a *P. stewartii* LrhA deletion mutant exhibits decreased surface motility and intermediate virulence levels in comparison to the wild type (14). However, little is known about the precise role of LrhA and its targets with regard to surface motility and virulence in *P. stewartii*.

In *Escherichia coli*, the function of LrhA is better understood. It is the key regulator controlling the expression of flagella, motility and chemotaxis by regulating the synthesis of $FlhD_2C_2$, the master regulator of flagella and chemotaxis gene expression (18). In *E. coli*, LrhA directly activates its own expression and represses the expression of *flhD/flhC*, thereby suppressing motility and chemotaxis (18). LrhA also controls *fimA* expression (19) and regulates RpoS translation (20, 21), but its binding site is not well defined (18).

In contrast to *E. coli*, *P. stewartii* is only capable of swarming rather than swimming motility. The bacterium's swarming motility is controlled by QS and contributes to its pathogenicity (22). The surface motility is flagellar-dependent since deletion of $fliC_I$ renders the bacterium aflagellar and incapable of swarming (22). There is no evidence demonstrating that EsaR plays a direct role in regulating flagella synthesis. However, EsaR does directly regulate *lrhA* in *P. stewartii* and thereby indirectly regulates surface motility and plant virulence (14) through unknown mechanisms. A transcriptome-level analysis of the LrhA regulon in *P. stewartii* showed that LrhA activates three genes and represses 23 genes four-fold or more (14).

In the present study, *CKS_0458* and *CKS_5211*, genes putatively encoding a fimbrial subunit and biosurfactant production enzyme, respectively, have now been confirmed to be direct targets of LrhA. In addition, LrhA has also been demonstrated to repress its own gene and that of RcsA. Follow-up studies led to the finding that RcsA also plays a role in surface motility. This work has helped further reveal how the QS regulatory cascade in *P. stewartii* coordinately controls genes important for interactions with the plant host.

MATERIALS AND METHODS

Strains and growth conditions.

Table 4.1 lists strains and plasmids used in this study. *E. coli* strains were grown in Luria-Bertani (LB) (10 g/l tryptone, 5 g/l yeast extract, and 5 g/l NaCl) broth or plates with 1.5% agar at 37°C. *P. stewartii* strains were grown in either LB or Rich Minimal (RM) medium (1X M9 salts, 2% casamino acids, 1 mM $MgCl_2$, and 0.4% glucose) at 30°C. Growth medium was supplemented with the following antibiotics: ampicillin (Ap, 100 μg/ml), chloramphenicol (Cm, 35 μg/ml), kanamycin (Kn, 50 μg/ml), nalidixic acid (Nal, 30 μg/ml), or streptomycin (Str, 100 μg/ml) as required (see Table 4.1).

Green fluorescent protein fusion (GFP) construction and testing.

A transcriptional fusion between the *lrhA* promoter (903 bp) and the gene for GFP was created through traditional molecular techniques as described previously (14). PCR primers (Table B.1) with the restriction sites *Eco*RI and *Kpn*I added to the 5' and 3' ends of the promoter sequence, respectively, were used to facilitate subcloning into the pPROBE'-GFP[tagless] vector (23). *E. coli* S17-1 was transformed with this plasmid construct containing P_{lrhA}-*gfp*, which was then moved into the wild-type *P. stewartii* DC283, *ΔlrhA* and *ΔlrhA/lrhA*$^+$ strains (Table B.1) via conjugation. The transconjugates were grown in RM medium supplemented with Kn and Nal

overnight, diluted in fresh medium to an OD_{600} of 0.05 at 30°C with shaking at 250 rpm to an OD_{600} of 0.2 to 0.5, diluted a second time in fresh medium to an OD_{600} of 0.025 and grown to an OD_{600} of 0.5. GFP measurements were done as previously described (14) with average relative fluorescence/OD_{600} from three experiments of triplicate samples, standard errors, and two-tailed homoscedastic Student's t-test values calculated for each strain.

Overexpression of LrhA.

The *lrhA* coding sequence was amplified using primers with *Bam*HI and *Hin*dIII sites (Table B.1), cloned into pGEM-T (Promega), and sequenced. After double digestion with *Bam*HI and *Hin*dIII, the construct was ligated into pET28a (Novagen) and transformed into *E. coli* (BL21-DE3) (24) to express LrhA with a His_6 tag at the N-terminus (37 kDa). Induction of protein expression with 0.1 M isopropyl β-D-1-thiogalactopyranoside (IPTG) was performed at an OD_{600} of 0.5 to 0.8, 19°C, overnight, shaking at 250 rpm. Cells were pelleted by centrifugation at 5,000 rpm in a JA-10 rotor (Beckman Coulter) for 20 min at 4°C, snap-frozen with liquid nitrogen and stored at -75°C. The cell pellet was then resuspended in Ni-NTA wash buffer (50 mM Tris-HCl, 300 mM NaCl, 50 mM imidazole) and sonicated to release proteins. Ultracentrifugation at 40,000 rpm in a Beckman Ti70 rotor for 1 h at 4°C was used to subsequently remove the cell debris. The protein was purified using a Ni-NTA column (HisTrap HP, GE Healthcare) with Ni-NTA elution buffer (50 mM Tris-HCl, 300 mM NaCl, 500 mM imidazole). The protein purity was observed through standard SDS-PAGE electrophoresis.

Electrophoretic mobility shift assays (EMSA).

Promoter regions of genes of interest were amplified with FAM-labeled primers (Table B.1) and extracted from a 1% agarose gel to examine the specific binding with purified His_6-LrhA over a range of concentrations. Twenty µl reactions with purified His_6-LrhA, 5 nM FAM-

DNA in 1X EMSA buffer (10% glycerol, 1 mM $MgCl_2$, 0.5 mM EDTA, 0.5 mM DTT, 50 mM NaCl, 10 mM Tris-HCl, 50 µg/ml poly (dI-dC) and 150 µg/ml BSA) were incubated at room temperature for 1 h before loading on to 1X TBE (10.8 g/l Tris-HCl, 5.5 g/l boric acid, 2 mM EDTA, pH 8.0) 4%, 5%, or 6% acrylamide native gels followed by electrophoresis at 80 V for 2 to 3 h. Images were visualized on a Typhoon Trio Scanner (GE Healthcare). Experiments were done in duplicate.

Construction of unmarked deletion mutant strains.

Chromosomal deletions of *CKS_0458/CKS_0459*, *CKS_5208*, *CKS_5211*, and *CKS_5211/CKS_5208* were constructed based on the Gateway system (Life Technologies) and suicide vectors as described previously (14), but with primers listed in Table B.1. In addition, another chromosomal deletion of *rcsA* was constructed using the same approach as described in a previous study (14), due to a deletion of ~66-kilobases (kb) in the chromosome of the original construct.

Construction of chromosomal complementation strains.

Complementation strains were constructed by generating a chromosomal insertion of the promoter and coding regions of the target gene into the neutral region downstream of *glmS* on the *P. stewartii* chromosome using the pUC18R6K-mini-Tn7-cat vector system (25) as previously described (14), but with primers listed in Table B.1.

Phenotypic surface motility assay.

Swarming motility for the wild-type, deletion and complementation strains was investigated under strict conditions to ensure a reproducible phenotype as previously described (14). Briefly, five µl of cell culture at an OD_{600} of 0.5 were spotted directly on the agar surface of

LB 0.4% agar quadrant plates supplemented with 0.4% glucose (22). Plates were incubated in a closed plastic box at 30^{o}C for 2 days prior to observation.

Phenotypic capsule production assay.

Bacterial strains were grown overnight in LB broth supplemented with the appropriate antibiotics at 30^{o}C with shaking. The overnight cultures were subcultured in fresh LB medium to an OD_{600} of 0.05 and grown to an OD_{600} of 0.5 at 30^{o}C with shaking. The strains were then cross-streaked with sterilized wooden sticks on 1.5% agar plates containing 0.1% casamino acids, 1% peptone and 1% glucose (CPG) (8, 14). Plates were incubated at 30^{o}C, lid-up for 2 days to observe the capsule production and visualized using the Bio-Rad Gel Doc imager system.

Plant virulence assay.

Virulence assays with *P. stewartii* strains in *Zea mays* seedlings were adapted from established methods (8) (14) with some modifications. In this study, *Zea mays* cv. Jubilee, 2B seeds were planted in Sunshine mix #1 or Promix soil for seven or six days, respectively, in a 28^{o}C growth chamber with ~100-200 μE m^{-2} s^{-1} light intensity, 16 h light/8 h dark and ~80% relative humidity (Percival Scientific, Inc.). Fifteen seedlings between 6-10 cm of height with two separated leaves were inoculated with five μl (~3 x 10^5 CFU) of bacterial culture grown to an OD_{600} of 0.2 in LB broth (~6 x 10^7 CFU/ml). Prior to plant inoculation, the bacterial cells were washed and resuspended in phosphate buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4 and 2 mM KH_2PO_4, pH 7.4). Wild-type strain and PBS controls were included in each trial, then, accumulated numbers of control-inoculated plants across all experiments were analyzed. A sterile needle (26G 5/8, 15.9 mm, SUB-Q Becton, Dickinson and Company) was used to make an ~1 cm incision in the stem ~1 cm above the soil line and the bacteria were added to the plant by slowly pipetting the inoculum while moving across the wound five times.

The plants were observed on day 12 post-infection to assess the virulence by two independent observers. Symptom severity was scored based on a five-point scale with 0 = no symptoms; 1 = few scattered lesions; 2 = scattered water soaking symptoms; 3 = numerous lesions and slight wilting; 4 = moderately severe wilt; 5 = death. The data for each treatment were averaged together and used to calculate the mean and standard error. A Student's t-test was used to calculate the p-value for experimental treatments compared to the wild-type treatment.

RESULTS

LrhA autorepresses its own gene expression in *P. stewartii*.

A GFP reporter was used to measure levels of transcription from the *lrhA* promoter in the wild-type, Δ*lrhA* and Δ*lrhA/lrhA*$^+$ strains of *P. stewartii* DC283 (Table 4.1). Expression levels of GFP in the Δ*lrhA* strain were significantly higher than the wild-type strain (Fig. 4.1, p = 0.00001) indicating that LrhA normally represses its own expression in the wild-type strain. The expression level of the *lrhA* promoter in the complementation Δ*lrhA/lrhA*$^+$ strain was restored to levels closer to those of the wild-type strain, and was also significantly different than the deletion strain (Fig. 4.1, p = 0.00002).

Identification of LrhA direct targets through EMSAs.

To determine if the observed *lrhA* autorepression occurred directly or indirectly, electrophoretic mobility shift assays (EMSAs) were performed. First, direct binding of LrhA to the promoter of its own gene was demonstrated by EMSA analysis (Fig. 4.2A). Next, the ability of LrhA to directly regulate additional gene targets was explored, using the *lrhA* promoter as a positive control for the His_6-LrhA activity and unlabeled P_{lrhA} DNA to prove the specificity of the binding. In *E. coli*, LrhA is known as a repressor of motility by direct interaction with the promoter region of *flhD/flhC*, whose products promote the expression of flagellar gene synthesis

(18). However, RNA-Seq data of expression levels of *flhD/flhC* in *P. stewartii* showed less than a two-fold difference between wild-type and Δ*lrhA* strains (14) suggesting a lack of transcriptional regulation. Here, EMSA analysis showed that His_6-LrhA does not bind to the promoter of *flhD/flhC* (Fig. 4.2B), explaining the observed lack of transcriptional regulation.

Additional analysis of the LrhA-regulated transcriptome in *P. stewartii* revealed that LrhA repressed the expression level of several more downstream targets, including *rcsA, CKS_0458*, *CKS_5208* and *CKS_5211* (14). RcsA activates capsule production, a known virulence factor in *P. stewartii* (11, 14, 26, 27). The putative roles of genes for fimbria encoded by *CKS_0458*, annotated as a putative fimbrial subunit, (and *CKS_0459* located downstream in an operon) and for surfactant expression encoded by *CKS_5208* and *CKS_5211*, annotated as a rhamnosyltransferase I subunit B and a putative alpha/beta superfamily hydrolase/acyltransferase, respectively, in plant colonization and/or virulence had not been established. However, it seemed plausible that they might also play roles in host association as they were some of the most highly LrhA-repressed genes, four-fold or greater (Kernell Burke et al, 2015). Therefore, the binding of LrhA to the promoters of these genes was also examined via EMSA. The direct binding of His_6-LrhA to P_{rcsA}, P_{CKS_0458} and P_{CKS_5211} (Fig. 4.2C-E), was demonstrated via EMSAs while P_{CKS_5208} did not interact with His_6-LrhA *in vitro* (Fig. 4.2F). Collectively, these findings identified four directly controlled gene targets in the LrhA regulon. The lack of LrhA regulation of $FlhD_2C_2$, the master regulator of flagellar-based motility and chemotaxis in *E. coli*, indicates a different role for LrhA in controlling *P. stewartii* motility. The direct binding of LrhA to the promoter of *rcsA* further links LrhA to *P. stewartii* pathogenesis. The role of the two other LrhA direct targets *CKS_0458* and *CKS_5211* remained to be established.

Examining the role of putative fimbrial and surfactant production genes in the surface motility and virulence of *P. stewartii*.

To further investigate the role of the downstream targets of LrhA putatively involved in production of fimbriae and surfactant, a reverse genetic approach was utilized. Markerless deletions of *CKS_0458/CKS_0459*, *CKS_5208*, *CKS_5211* and *CKS_5211*/*CKS_5208* were successfully generated. Corresponding chromosomal complementation strains were also generated with the exception of a double deletion mutant of *CKS_5211*/*CKS_5208* complementation strain, due to the length constraint of the DNA fragment containing the adjacent genes. In surface motility assays, the *P. stewartii* wild-type strain showed either uni-directional (Fig. 4.3A & B.1A) or omni-directional (Fig. 4.3B & B.1B) expansion from the inoculum sites as had been previously observed (14, 22). In comparison to the wild type, there is no obvious difference between the various deletion and complementation strains; they all possessed similar level of expansion on the agar surface (Fig. B.1C-J). Therefore, these genes do not appear to play any detectable role in surface motility via this assay.

The same deletion and complementation strains of the genes putatively involved in fimbriae and surfactant production were also tested for virulence via *in planta* xylem infection assays. A *lrhA* deletion strain caused an intermediate level of disease severity in corn seedlings during xylem-infection assays (14). However, similar to the surface motility assays, no significant impacts on the virulence of *P. stewartii* were observed in the strains with deletions in either the fimbriae or surfactant synthesis genes (Fig. B.2). Hence, the contribution of these genes individually to the virulence of the phytopathogen could not be measured.

Re-examining the role of RcsA in the capsule production, surface motility and virulence of *P. stewartii*.

The important finding that LrhA directly binds to the promoter of *rcsA*, led to a reexamination of the previous findings about the physiological role of RcsA in *P. stewartii*. In prior work, *rcsA* deletion and complementation strains of DC283 had been constructed (Δ*rcsA*-2015 and Δ*rcsA/rcsA*$^+$-2015) (14). However, complete assembly of the genome of *P. stewartii* DC283 (28) revealed that there is a large deletion, ~66 kb containing 68 genes (Table B.2), in the Δ*rcsA*-2015 and Δ*rcsA/rcsA*$^+$-2015 strains. This deletion was not obvious using the incomplete genome sequence (NCBI GenBank accession no. AHIE00000000.1), but was found during a re-analysis of previously generated RNA-Seq data (14) using the new genome sequence (NCBI GenBank accession no. CP017581).

Therefore, a new set of *rcsA* deletion and complementation strains was re-constructed (Δ*rcsA*-2017 and Δ*rcsA/rcsA*$^+$-2017) and shown to include the 66-kb region using PCR (data not shown). These new strains then were subjected to three assays to establish the true phenotypes of the *rcsA* deletion strain. First, capsule production assays have re-confirmed that RcsA regulates EPS production, as was shown for the 2015 strains (14). Both the Δ*rcsA*-2015 (Fig. 4.4D) and Δ*rcsA*-2017 strains (Fig. 4.4F) are not as mucoid as the parental wild-type strain (Fig. 4.4A) or the pair of *lrhA* deletion and complementation strains (Fig. 4.4B & 4.4C) as assessed by visual observation. The chromosomal complementation strains Δ*rcsA/rcsA*$^+$-2015 (Fig. 4.4E) and Δ*rcsA/rcsA*$^+$-2017 (Fig. 4.4G) had mucoid levels as high or higher than those seen in the wild type (Fig. 4.4A).

Second, surface motility assays were performed. These original strains Δ*rcsA*-2015 and Δ*rcsA/rcsA*$^+$-2015 strains had not previously been examined for surface motility ((14), but both

were surprisingly defective for this phenotype (Fig. 4.3E and 4.3F). The fact that surface motility was not complemented by addition of *rcsA* back into the chromosome provided further evidence for the importance of the 66-kb region deletion that had been discovered initially through bioinformatics analysis. The new Δ*rcsA*-2017 also has severely reduced surface movement (Fig. 4.3G) while its complementation strain (Fig. 3H) restored motility levels similar to the wild-type strain (Fig. 4.3A and 4.3B). Thus, it has been demonstrated that RcsA plays a previously unappreciated role in the surface motility of *P. stewartii.* The defect in surface motility associated with the deletion of *rcsA* (Fig. 4.3G) appears to be greater than the defect in Δ*lrhA* (Fig. 4.3C). The Δ*lrhA/lrhA*$^{+}$ strain has restored levels of surface motility (Fig. 4.3D), similar to the wild type (Fig. 4.3A & 4.3B), as previously reported (14).

Finally, the xylem infection assays for the newly constructed Δ*rcsA*-2017 strain and its complement, with the inclusion of the wild-type strain and PBS as controls, indicated that the absence of *rcsA* significantly ($p < 0.05$) reduces the severity of the disease compared to the wild-type and complementation strains (Fig. 4.5) ($p < 0.05$). These results have similar trends with those reported for the 2015 strains (14) which confirms the role of *rcsA* in virulence of this phytopathogen. However, strains from the 2015 study that were missing the 66- kb region were reduced in their average disease severity (score ~0 and ~1.5 for the deletion and complementation strains, respectively) in comparison to the new 2017 strains (score ~1.5 and ~3.5 for the deletion and complementation strains, respectively) while the wild-type control had similar levels in both studies, implicating a role for the 66-kb region in virulence as well as surface motility.

DISCUSSION

The role of the LrhA regulon in *P. stewartii* was further investigated in this study to understand how it is involved in the surface motility and virulence of the pathogen. Previous studies showed that surface motility in *P. stewartii* contributes to disease pathogenesis and this process involves both QS-controlled biofilm formation and flagella (22). However, to date, there is no clear evidence to directly connect the synthesis of flagella to QS control in *P. stewartii*. Unlike *E. coli*, the QS-controlled transcription factor LrhA in *P. stewartii* does not regulate $FlhD_2C_2$, the master activator of flagellar synthesis. This was suggested by earlier RNA-Seq data (14), but directly tested here through EMSA that confirmed the inability of LrhA to bind to the *flhD/flhC* promoter. Additionally, LrhA activates its own expression in *E. coli* whereas autorepression was observed in *P. stewartii*. Even though *P. stewartii* LrhA has 77% amino acid identity to *E. coli* LrhA, the two have clearly evolved distinctive physiological roles in their host organisms.

In an attempt to define the function of the genes controlled by LrhA in *P. stewartii*, a reverse genetics approach was used to examine the role of select LrhA-regulated genes in surface motility and virulence of the phytopathogen. Multiple deletion and complementation strains of genes annotated as being involved in surfactant production (*CKS_5208* and *CKS_5211*, initially annotated as a rhamnosyltransferase I subunit B and putative alpha/beta superfamily hydrolase/acyltransferase, respectively) and fimbriae assembly (*CKS_0458* and *CKS_0459*, annotated as putative fimbrial subunits) were constructed and tested. Interestingly, none of these genes appear to play a fundamental role in surface motility and virulence individually. A LrhA deletion mutant impacting expression of multiple genes in the regulon produced noticeably

decreased surface motility, but only intermediate virulence levels in comparison to the wild-type strain (14).

With regard to biosurfactant and fimbriae genes potentially associated with surface motility and adhesion, respectively, *P. stewartii* appears to utilize multiple levels of repression to ensure that the level of those genes' expression is minimal. This low level of expression was again confirmed by an *in planta* RNA-Seq analysis (29). In the LrhA deletion strain expression of these genes was elevated. Thus, deletion mutants might actually mimic wild-type levels of the expression of these genes, producing a wild-type phenotype. Alternatively, these genes are not functional in the wild-type strain (indeed the new genome sequence (28) suggests that *CKS_5211* is a pseudogene) or they may serve another function for the bacterium that was not examined in this study. Biofilm/adhesion assays were inconclusive (data not shown). Interestingly, some *Pantoea* species have been demonstrated to produce biosurfactants when grown on hydrocarbons (30). How this might impact bacterial surface motility or survival *in planta* is unclear.

It has been demonstrated that both RcsA and LrhA play an essential role to the surface motility of the wild-type strain of *P. stewartii*. The observed intermediate impact of a LrhA deletion on virulence may be due primarily to its direct control of RcsA and thereby its indirect control on the levels of stewartan extracellular polysaccharide produced during growth within the plant. However, it could be that some of the other genes regulated by LrhA that were not examined in this work were actually contributing to the observed phenotypes in the LrhA deletion strain. RNA-Seq analysis of the transcriptome controlled by LrhA revealed 23 additional genes, in addition to the ones examined in this study, that were differentially expressed four-fold or more in comparison to the wild-type strain (14). Overall, the majority of the genes in the LrhA regulon code for hypothetical proteins and phage-related proteins, 57.7%

(15/26) and 15.4% (4/26) respectively. The possible role of these genes with regard to surface motility and virulence remains to be established, but LrhA clearly regulates these processes.

The newly discovered connection between RcsA and surface motility suggests coordination of the RcsA and LrhA regulons with regard to bacterial virulence in the corn host beyond promotion of capsule production. Capsule production is thought to be a factor impacting the ability of surface motility to occur in this phytopathogen, which may explain the need for integrated downstream regulation. The fact that the strain with the 66-kb deletion region could not be complemented by *rcsA* suggests that there are additional genes in this region that are essential to surface motility and virulence. Further work will be needed to identify these genes and to overall correlate to the ability of the phytopathogen to move inside the plant via surface motility in relation to virulence.

CONCLUSIONS

The findings of this study have further defined the tightly coordinated gene regulation that occurs in the QS regulon of the corn pathogen *P. stewartii.* The EsaR-activated transcription factor LrhA was found to directly auto-repress expression of its own gene as demonstrated through GFP-transcription fusions and EMSA experiments. In addition, the direct binding of LrhA to downstream targets, such as the promoters of genes coding for RcsA, and for putative biosurfactant synthesis (*CKS_5211*) and fimbrial production (*CKS_0458*), was also shown. This established a hierarchy of gene regulation in the QS network from the master regulator, EsaR, to the downstream transcription factors, RcsA and LrhA, which in turn control the expression of their own targets. Intriguingly, EsaR also directly controls some of these same targets (12, 13) integrating with coherent type two (RcsA) and type three (LrhA) feed forward loops (31) to

regulate genes in the QS regulon in a manner that ensures precisely synchronized gene expression (Fig. 4.6).

ACKNOWLEDGMENTS

We thank Roderick Jensen for his critical review of the manuscript. We thank the laboratories of Rich Helm and Brenda Winkel for sharing their plant growth chamber.

REFERENCES

1. **Esker PD, Nutter FW.** 2002. Assessing the risk of Stewart's disease of corn through improved knowledge of the role of the corn flea beetle vector. Phytopathology **92:**668-670.
2. **Ham JH, Majerczak DR, Arroyo-Rodriguez AS, Mackey DM, Coplin DL.** 2006. WtsE, an AvrE-family effector protein from *Pantoea stewartii* subsp. *stewartii*, causes disease-associated cell death in corn and requires a chaperone protein for stability. Molecular Plant-Microbe Interactions **19:**1092-1102.
3. **Braun EJ.** 1982. Ultrastructural investigation of resistant and susceptible maize inbreds infected with *Erwinia stewartii*. Phytopathology **72:**159-166.
4. **Roper MC.** 2011. *Pantoea stewartii* subsp. *stewartii*: lessons learned from a xylem-dwelling pathogen of sweet corn. Molecular Plant Pathology **12:**628-637.
5. **von Bodman SB, Bauer WD, Coplin DL.** 2003. Quorum sensing in plant-pathogenic bacteria. Annual Review of Phytopathology **41:**455-482.
6. **Beck von Bodman S, Farrand SK.** 1995. Capsular polysaccharide biosynthesis and pathogenicity in *Erwinia stewartii* require induction by an N-acylhomoserine lactone autoinducer. Journal of Bacteriology **177:**5000-5008.
7. **von Bodman SB, Ball JK, Faini MA, Herrera CM, Minogue TD, Urbanowski ML, Stevens AM.** 2003. The quorum sensing negative regulators EsaR and ExpR(Ecc), homologues within the LuxR family, retain the ability to function as activators of transcription. Journal of Bacteriology **185:**7001-7007.
8. **von Bodman SB, Majerczak DR, Coplin DL.** 1998. A negative regulator mediates quorum-sensing control of exopolysaccharide production in *Pantoea stewartii* subsp. *stewartii*. Proceedings of the National Academy of Sciences of the United States of America **95:**7687-7692.
9. **Shong J, Huang YM, Bystroff C, Collins CH.** 2013. Directed evolution of the quorum-sensing regulator EsaR for increased signal sensitivity. ACS Chemical Biology **8:**789-795.
10. **Schu DJ, Carlier AL, Jamison KP, von Bodman S, Stevens AM.** 2009. Structure/function analysis of the *Pantoea stewartii* quorum-sensing regulator EsaR as an activator of transcription. Journal of Bacteriology **191:**7402-7409.
11. **Minogue TD, Carlier AL, Koutsoudis MD, von Bodman SB.** 2005. The cell density-dependent expression of stewartan exopolysaccharide in *Pantoea stewartii* ssp. *stewartii* is a function of EsaR-mediated repression of the *rcsA* gene. Molecular Microbiology **56:**189-203.
12. **Ramachandran R, Stevens AM.** 2013. Proteomic analysis of the quorum-sensing regulon in *Pantoea stewartii* and identification of direct targets of EsaR. Applied and Environmental Microbiology **79:**6244-6252.
13. **Ramachandran R, Burke AK, Cormier G, Jensen RV, Stevens AM.** 2014. Transcriptome-based analysis of the *Pantoea stewartii* quorum-sensing regulon and identification of EsaR direct targets. Applied and Environmental Microbiology **80:**5790-5800.
14. **Kernell Burke A, Duong DA, Jensen RV, Stevens AM.** 2015. Analyzing the transcriptomes of two quorum-sensing controlled transcription factors, RcsA and LrhA, important for *Pantoea stewartii* virulence. PLOS ONE **10:**e0145358.
15. **Carlier AL, von Bodman SB.** 2006. The *rcsA* promoter of *Pantoea stewartii* subsp. *stewartii* features a low-level constitutive promoter and an EsaR quorum-sensing-regulated promoter. Journal of Bacteriology **188:**4581-4584.
16. **Nimtz M, Mort A, Wray V, Domke T, Zhang Y, Coplin DL, Geider K.** 1996. Structure of stewartan, the capsular exopolysaccharide from the corn pathogen *Erwinia stewartii*. Carbohydrate Research **288:**189-201.

17. **Carlier A, Burbank L, von Bodman SB.** 2009. Identification and characterization of three novel EsaI/EsaR quorum-sensing controlled stewartan exopolysaccharide biosynthetic genes in *Pantoea stewartii* ssp. *stewartii*. Molecular Microbiology **74:**903-913.
18. **Lehnen D, Blumer C, Polen T, Wackwitz B, Wendisch VF, Unden G.** 2002. LrhA as a new transcriptional key regulator of flagella, motility and chemotaxis genes in *Escherichia coli*. Molecular Microbiology **45:**521-532.
19. **Blumer C, Kleefeld A, Lehnen D, Heintz M, Dobrindt U, Nagy G, Michaelis K, Emody L, Polen T, Rachel R, Wendisch VF, Unden G.** 2005. Regulation of type 1 fimbriae synthesis and biofilm formation by the transcriptional regulator LrhA of *Escherichia coli*. Microbiology **151:**3287-3298.
20. **Peterson CN, Carabetta VJ, Chowdhury T, Silhavy TJ.** 2006. LrhA regulates *rpoS* translation in response to the *Rcs* phosphorelay system in *Escherichia coli*. Journal of Bacteriology **188:**3175-3181.
21. **Gibson KE, Silhavy TJ.** 1999. The LysR homolog LrhA promotes RpoS degradation by modulating activity of the response regulator *sprE*. Journal of Bacteriology **181:**563-571.
22. **Herrera CM, Koutsoudis MD, Wang X, von Bodman SB.** 2008. *Pantoea stewartii* subsp. *stewartii* exhibits surface motility, which is a critical aspect of Stewart's wilt disease development on maize. Molecular Plant-Microbe Interactions **21:**1359-1370.
23. **Miller WG, Leveau JH, Lindow SE.** 2000. Improved *gfp* and *inaZ* broad-host-range promoter-probe vectors. Molecular Plant-Microbe Interactions **13:**1243-1250.
24. **Studier FW, Moffatt BA.** 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. Journal of Molecular Biology **189:**113-130.
25. **Choi KH, Gaynor JB, White KG, Lopez C, Bosio CM, Karkhoff-Schweizer RR, Schweizer HP.** 2005. A Tn7-based broad-range bacterial cloning and expression system. Nature methods **2:**443-448.
26. **Poetter K, Coplin DL.** 1991. Structural and functional analysis of the *rcsA* gene from *Erwinia stewartii*. Molecular & General Genetics **229:**155-160.
27. **Wehland M, Kiecker C, Coplin DL, Kelm O, Saenger W, Bernhard F.** 1999. Identification of an RcsA/RcsB recognition motif in the promoters of exopolysaccharide biosynthetic operons from *Erwinia amylovora* and *Pantoea stewartii* subspecies *stewartii*. The Journal of Biological Chemistry **274:**3300-3307.
28. **Duong DA, Stevens AM, Jensen RV.** 2017. Complete genome assembly of *Pantoea stewartii* subsp. *stewartii* DC283, a corn pathogen. Genome Announcements **5**: e00435-17.
29. **Packard H, Kernell Burke A, Jensen RV, Stevens AM.** 2017. Analysis of the *in planta* transcriptome expressed by the corn pathogen *Pantoea stewartii* subsp. *stewartii* via RNA-Seq. PeerJ **5:**e3237.
30. **Vasileva-Tonkova E, Gesheva V.** 2007. Biosurfactant production by antarctic facultative anaerobe *Pantoea* sp. during growth on hydrocarbons. Current Microbiology **54:**136-141.
31. **Mangan S, Alon U.** 2003. Structure and function of the feed-forward loop network motif. Proceedings of the National Academy of Sciences of the United States of America **100:**11980-11985.
32. **Dolph PJ, Majerczak DR, Coplin DL.** 1988. Characterization of a gene cluster for exopolysaccharide biosynthesis and virulence in *Erwinia stewartii*. Journal of Bacteriology **170:**865-871.
33. **Grant SG, Jessee J, Bloom FR, Hanahan D.** 1990. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. Proceedings of the National Academy of Sciences of the United States of America **87:**4645-4649.
34. **Kvitko BH, Bruckbauer S, Prucha J, McMillan I, Breland EJ, Lehman S, Mladinich K, Choi KH, Karkhoff-Schweizer R, Schweizer HP.** 2012. A simple method for construction of *pir*+ Enterobacterial hosts for maintenance of R6K replicon plasmids. BMC Research Notes **5:**157.

35. **Simon R, Priefer U, Pühler A.** 1983. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in Gram negative bacteria. Nature Biotechnology **1:**784 - 791.
36. **Labes M, Puhler A, Simon R.** 1990. A new family of RSF1010-derived expression and *lac*-fusion broad-host-range vectors for gram-negative bacteria. Gene **89:**37-46.
37. **Stabb EV, Ruby EG.** 2002. RP4-based plasmids for conjugation between *Escherichia coli* and members of the *Vibrionaceae*. Methods in Enzymology **358:**413-426.

Table 4.1: Strains and plasmids used in the study.

Strains	Genotype and notes[a]	References
***Pantoea stewartii* strains**		
DC283	Wild-type strain; Nal^r	(32)
Δ*lrhA*	Unmarked deletion of *lrhA* coding sequence from DC283; Nal^r	(14)
Δ*lrhA/lrhA*$^+$	Δ*lrhA* with chromosomal complementation of *lrhA* and its promoter downstream of *glmS*; Nal^r Cm^r	(14)
Δ*rcsA*-2015	Unmarked deletion of *rcsA* coding sequence from DC283; Nal^r, missing 66-kb region	(14)
Δ*rcsA/rcsA*$^+$-2015	Δ*rcsA* with chromosomal complementation of *rcsA* and its promoter downstream of *glmS*; Nal^r Cm^r, missing 66-kb region	(14)
Δ*rcsA*-2017	Unmarked deletion of *rcsA* coding sequence from DC283; Nal^r	This study
Δ*rcsA/rcsA*$^+$-2017	Δ*rcsA* with chromosomal complementation of *rcsA* and its promoter downstream of *glmS*; Nal^r Cm^r	This study
Δ*CKS_0458-CKS_0459*	Unmarked deletion of both *CKS_0458* and *CKS_0459* coding sequence from DC283; Nal^r	This study
Δ*CKS_0458-CKS_0459 /CKS_0458*$^+$	Δ*CKS_0458-CKS_0459* with chromosomal complementation of *CKS_0458* and its promoter downstream of *glmS*; Nal^r Cm^r	This study
Δ*CKS_0458-CKS_0459 /CKS_0458-CKS_0459*$^+$	Δ*CKS_0458-CKS_0459* with chromosomal complementation of *CKS_0458-CKS_0459* and their promoter downstream of *glmS*; Nal^r Cm^r	This study
Δ*CKS_5208*	Unmarked deletion of *CKS_5208* coding sequence from DC283; Nal^r	This study
Δ*CKS_5208/CKS_5208*$^+$	Δ*CKS_5208* with chromosomal complementation of *CKS_5208* and its promoter downstream of *glmS*; Nal^r Cm^r	This study
Δ*CKS_5211*	Unmarked deletion of *CKS_5211* coding sequence from DC283; Nal^r	This study
Δ*CKS_5211/CKS_5211*$^+$	Δ*CKS_5211* with chromosomal complementation of *CKS_5211* and its promoter downstream of *glmS*; Nal^r Cm^r	This study
Δ*CKS_5211/*Δ*CKS_5208*	Unmarked deletion of both *CKS_5211* and *CKS_5208* coding sequences from DC283; Nal^r	This study
***Escherichia coli* strains**		
Top 10	*F*$^-$ *mcrA* Δ(*mrr-hsdRMS-mcrBC*) *Φ80dlacZΔM15* Δ*lacX*74 *deoR recAI araD139* Δ(*ara-leu)7697 galU galK rpsL* (Str^r) *endA1 nupG*	(33)

DH5α *λpir*	*F- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169 hsdR17(rK-mK+) λpir*	(34)
S17-1	*recA pro hsdR RP4-2-Tc::Mu-Km::Tn7*	(35)
S17-1 *λpir*	*recA pro hsdR RP4-2-Tc::Mu-Km::Tn7 λpir*	(36)
BL21-DE3	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*	(24)
Plasmids		
pGEM-T	Cloning vector, Ap^r	Promega
pET28a	Expression vector, Kn^r	Novagen
pDONR201	Entry vector in the Gateway system, Kn^r	Life Technologies
pAUC40	Suicide vector pKNG101::*attR-ccdB*-Cm^R; Cm^r, Str^r, *sacB*	(17)
pEVS104	Conjugative helper plasmid, *tra trb;* Kn^r	(37)
pUC18R6K-mini-Tn7-cat	Tn7 vector for chromosomal integration into the intergenic region downstream of *glmS*; Cm^r, Ap^r	(25)
pPROBE'-GFP[tagless] P_{lrhA}	pPROBE'-GFP[tagless] vector with the promoter of *lrhA;* Kn^r	This study

[a] Ap^r, ampicillin resistance; Nal^r, nalidixic acid resistance; Kn^r, kanamacyin resistance; Cm^r, chloramphenicol resistance; Str^r, streptomycin resistance.

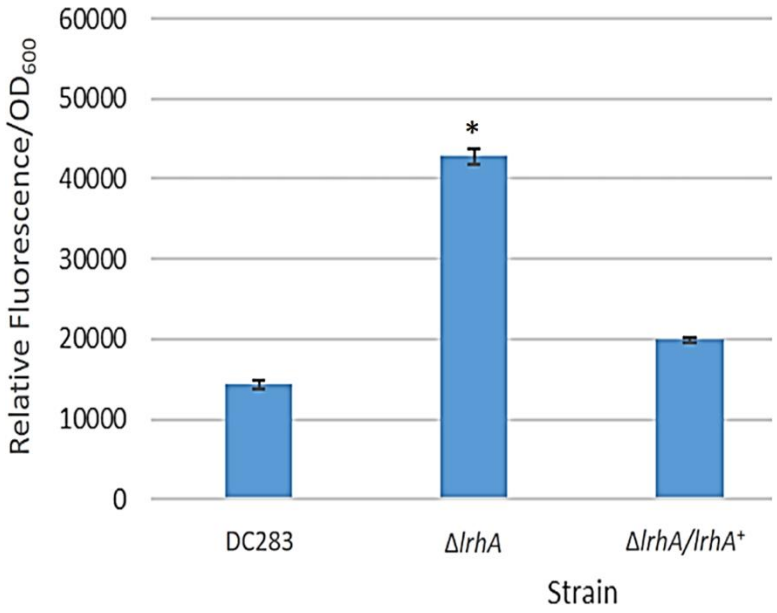

Figure 4.1. Expression levels of a *lrhA* promoter-*gfp* transcription reporter in three *P. stewartii* strains. The wild-type DC283 and Δ*lrhA* and Δ*lrhA/lrhA*$^+$ strains in the same genetic background (containing pPROBE'-GFP[tagless] P_{lrhA}) were grown to an OD_{600} of 0.5 and GFP expression levels from the *lrhA* promoter-*gfp* transcription reporter were measured as average relative fluorescence/OD_{600}. Data represents three experimental samples analyzed in triplicate. Error bars denote standard error. The asterisk (*) represents a statistically significant difference ($p<0.05$) between the Δ*lrhA* and both the wild-type and Δ*lrhA/lrhA*$^+$ strains using a two-tailed homoscedastic Student's t-test.

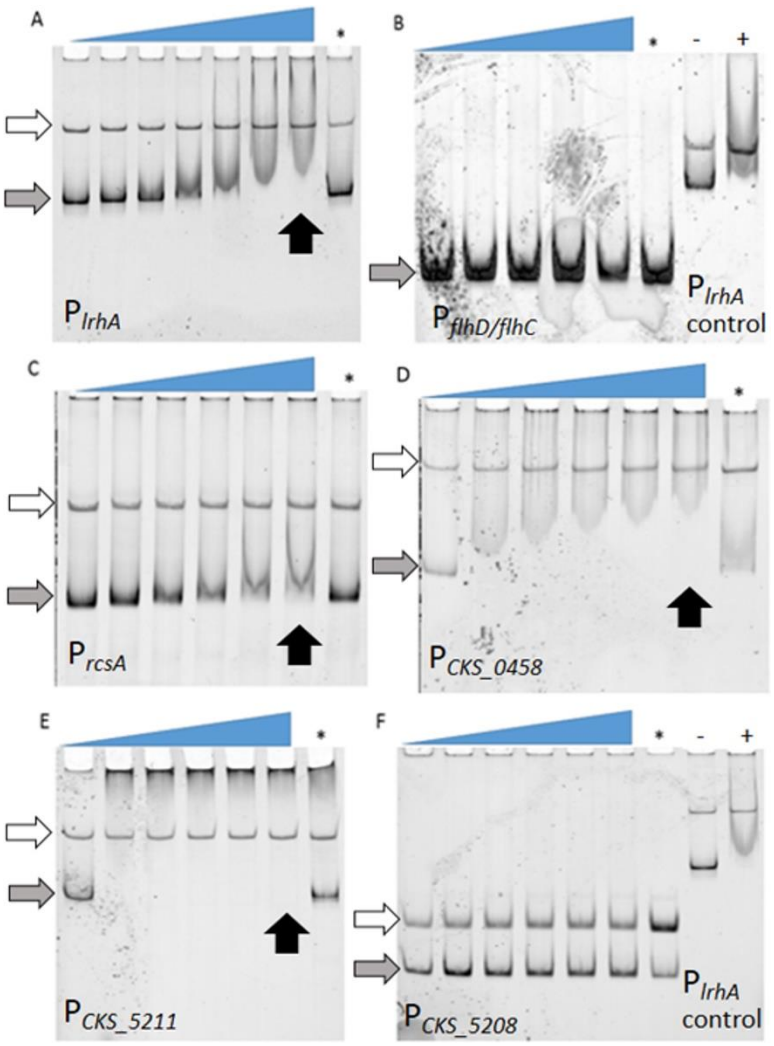

Figure 4.2. Examination of binding of LrhA to select target promoters via EMSA. FAM-DNA probes were incubated with increasing concentrations of His_6-LrhA (LrhA) from left to right, corresponding to the slope of the triangles, to investigate the mobility shift upon specific binding to the protein. The competition reaction (indicated by the asterisk, *) was conducted with 25 nM unlabeled DNA of P_{lrhA} to prove the specificity of the interaction. Autoregulation of LrhA was confirmed with the direct binding between purified LrhA to its promoter (panel A). Shifted bands were also observed with P_{rcsA} (panel C), P_{CKS_0458} (panel D), and P_{CKS_5211} (panel E). There were no shifted bands observed for P_{flhDC} (panel B) and P_{CKS_5208} (panel F), while the positive controls for LrhA activity showed a shift (-: reaction with P_{lrhA} probe in the absence of LrhA, +:

reaction with P_{lrhA} probe in the presence of 200 nM LrhA). Concentrations of LrhA tested for P_{lrhA} (panel A) are 0, 25, 50, 100, 200, 400, and 800 nM. Concentrations of LrhA tested for P_{flhDC} (panel B) are 0, 400, 600, 800, and 1000 nM. Concentrations of LrhA tested for P_{rcsA} (panel C), P_{CKS_0458} (panel D), P_{CKS_5211} (panel E) and P_{CKS_5208} (panel F) are 0, 200, 400, 600, 800, and 1000 nM. Grey arrows highlight unbound DNA probes. White arrows indicate unbound DNA generated during PCR reactions that do not interact specifically with LrhA. Black arrows point to the lane with specific binding at the highest concentration of LrhA.

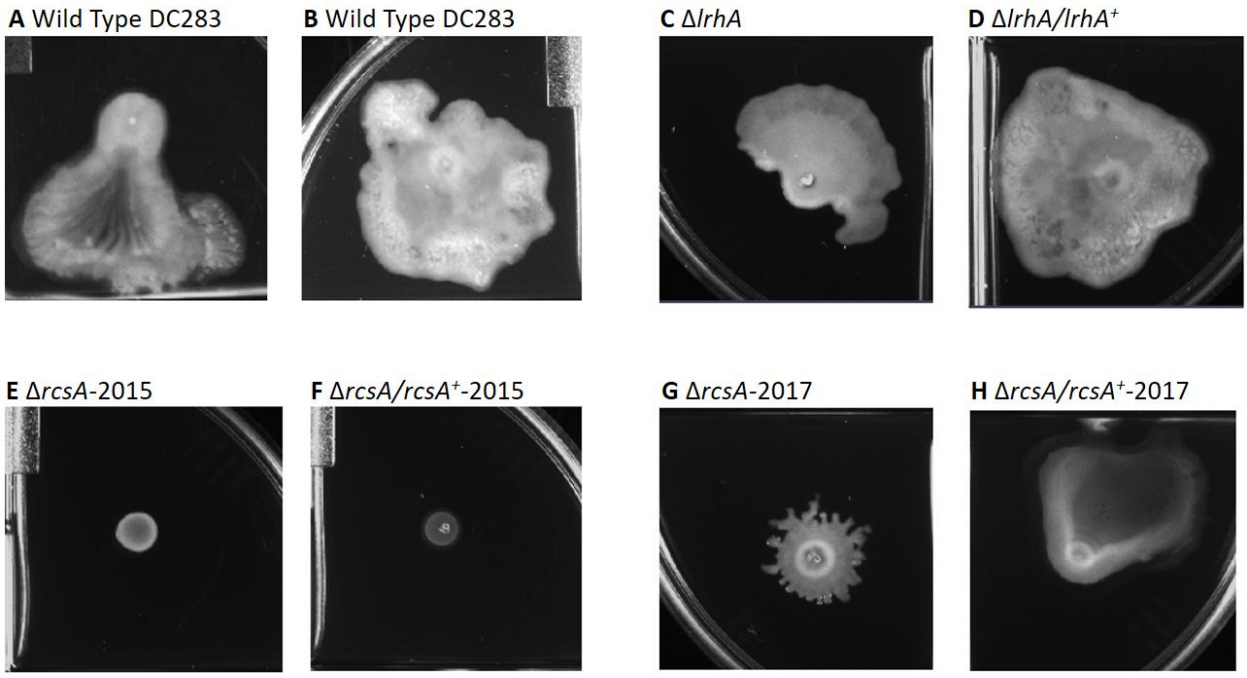

Figure 4.3. Impact of RcsA and LrhA on surface motility of *P. stewartii*. The pictures show the analysis of surface motility in *P. stewartii* DC283 strains. Examples of wild type unidirectional (panel A) or omnidirectional surface motility (panel B) are shown as controls. The Δ*lrhA/lrhA*$^+$ complementation strain (panel D) is similar to the control in panel B, while the Δ*lrhA* strain has reduced surface motility expanding over a smaller surface area (panel C), as has been previously observed (14). Both Δ*rcsA* strains had dramatically reduced surface motility (panels E and G) as well as the Δ*rcsA/rcsA*$^+$-2015 strain (panel F). The Δ*rcsA/rcsA*$^+$-2017 strain was complemented for the defect in surface motility (panel H). All pictures were taken at the same magnification after 2 days of incubation at 30^{o}C in a closed plastic box.

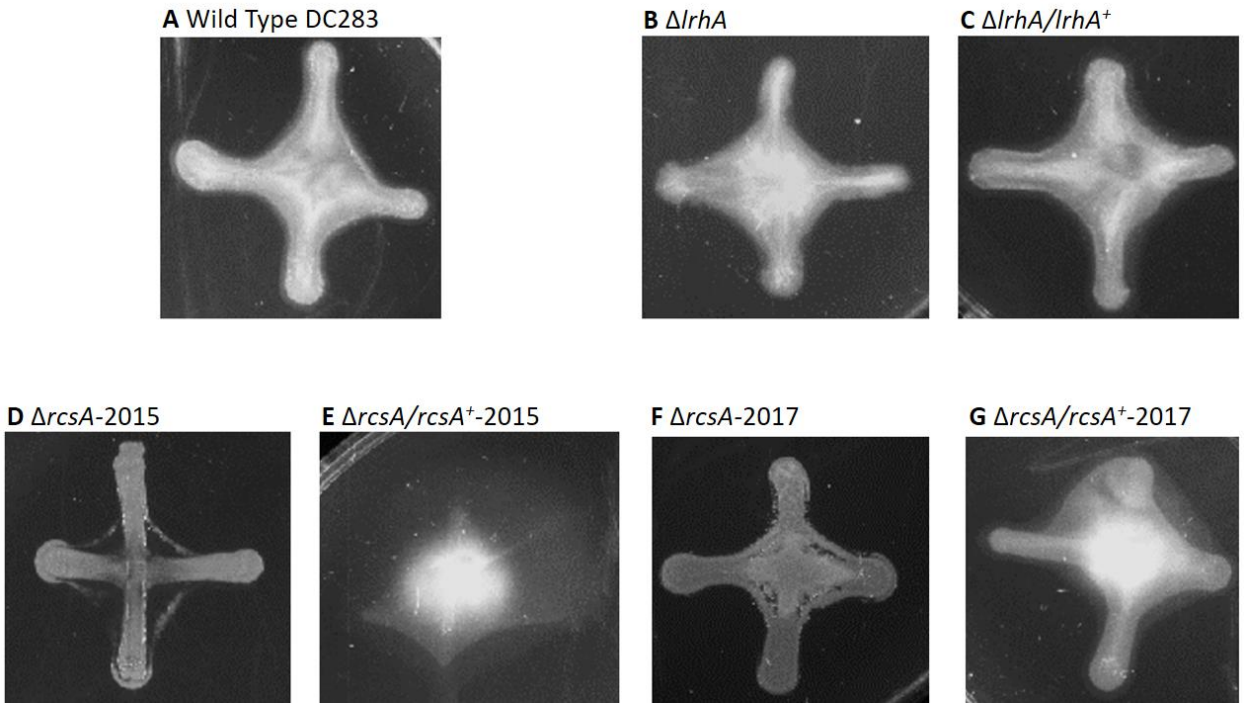

Figure 4.4. Impact of RcsA and LrhA on capsule production of *P. stewartii*. All pictures were taken at the same magnification after two days of incubation at 30^{o}C after cross-streaking on casamino acid, peptone, glucose (CPG) agar plates. Differences in capsule production are apparent in the regions between the arms of the X-cross streak.

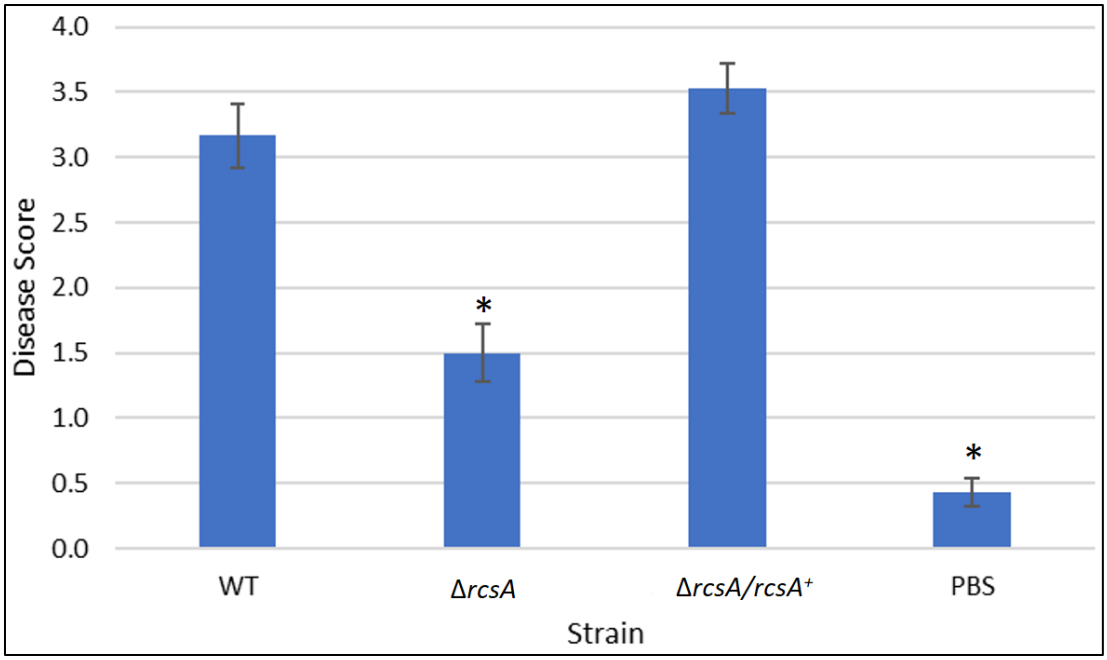

Figure 4.5. Plant assay testing the role of RcsA in virulence. Data shown is the average score of disease for Day 12 of an infection assay performed with 15 plants inoculated with *P. stewartii* DC283 strains: wild type (WT), Δ*rcsA*-2017, Δ*rcsA/rcsA*$^{+}$-2017, or PBS as a negative control. Higher value in the disease score indicates more severe symptoms from the infection. The asterisks (*) represent strains that are statistically significantly different ($p < 0.05$) from the wild-type strain using a two-tailed homoscedastic Student's t-test. Error bars denote standard error.

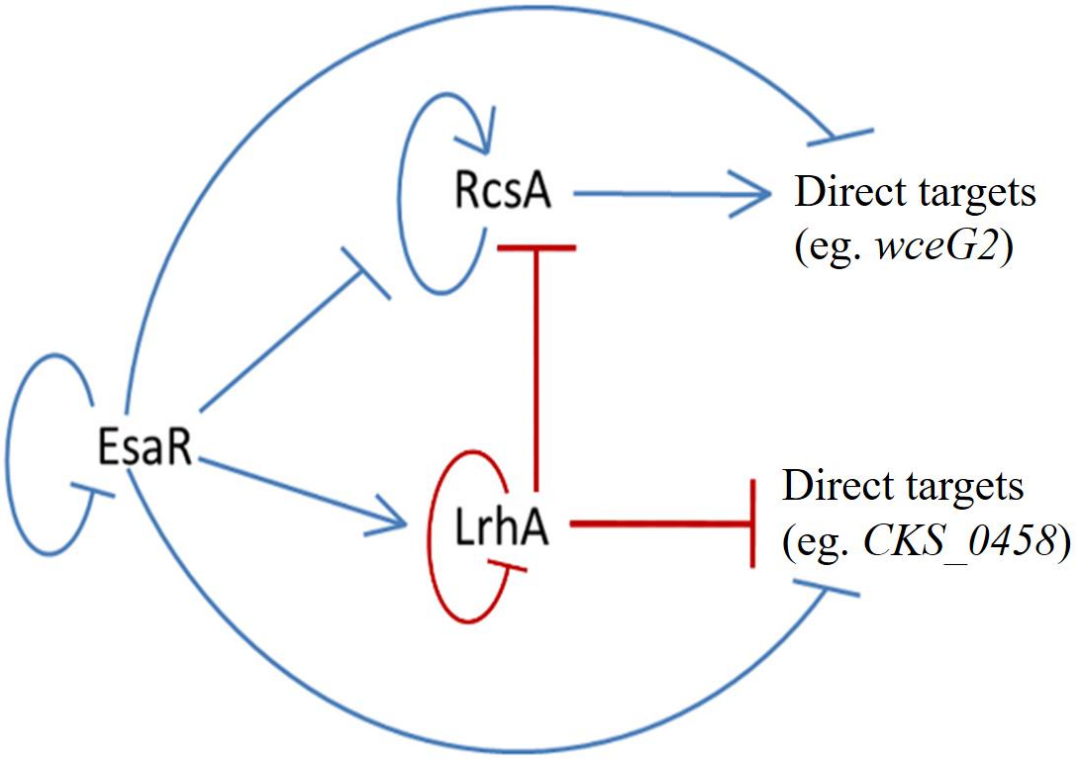

Figure 4.6. Updated model of the quorum-sensing regulatory network in *P. stewartii*. Solid lines indicate known direct regulatory control. Red lines indicate direct control found in this study. Arrows represent activation and T lines represent repression. At low cell density when AHL levels are low, EsaR represses expression of *rcsA, wceG2*, and *CKS_0458*, and activates expression of *lrhA*. LrhA represses its own expression as well as that of *rcsA* and *CKS_0458*. At high cell density when EsaR-AHL complexes form, EsaR no longer activates or represses its direct targets. Thus, *rcsA* expression increases leading to activation of *wceG2* and other genes necessary for capsule production. See the text for additional details.

Chapter Five

Understanding Plant-microbe Interactions by Identification of *Pantoea stewartii* subsp. *stewartii* Genes Important for Survival in Corn Xylem

Duy An Duong, Roderick V. Jensen, Ann M. Stevens#. In preparation. Understanding plant-microbe interactions by identification of *Pantoea stewartii* subsp. *stewartii* genes important for survival in corn xylem.

Address correspondence to Ann M. Stevens, ams@vt.edu

Running Head: *Pantoea stewartii* genes important for plant survival
Key words: *Pantoea stewartii*, phytopathogen, Tn-Seq; wilt disease; xylem dwelling

ABSTRACT

The bacterium *Pantoea stewartii* subsp. *stewartii* causes Stewart's wilt disease in corn. *P. stewartii* is transmitted to plants via corn flea beetles where it first colonizes the apoplast, then migrates to the xylem and forms a biofilm. This blocks water transport leading to wilt disease. Quorum sensing insures that capsule production occurs only at high cell density. Identifying additional bacterial genes essential for survival *in planta* will provide insights into the plant-microbe interactions occurring during wilt disease. The complete *P. stewartii* genome assembly facilitated a genomic-level analysis using a Tn-Seq approach. A mariner transposon library of approximately 40,000 mutants was constructed and used to inoculate corn seedlings through a xylem infection model, as well as two *in vitro* growth conditions, LB and M9 minimal media. Tn-Seq analysis showed that the number of transposon mutations are reduced more than ten-fold for 377 genes *in planta* compared to the library that grew in LB, suggesting they are important for *in planta* survival. Interestingly, a small set of genes had a higher abundance of mutants *in planta* versus *in vitro* conditions, suggesting enhanced strain fitness with loss of those genes inside the host. *In planta* competition assays retested the trends of the Tn-Seq data for several genes including two outer membrane proteins, Lon protease and two quorum-sensing associated transcription factors, RcsA and LrhA. Virulence assays were performed to check for correlation between growth/colonization and pathogenicity. This study demonstrates the capacity of a Tn-Seq approach to advance our understanding of *P. stewartii*-corn interactions.

INTRODUCTION

The genus *Pantoea* is a diverse group in the *Enterobacteriaceae* family comprised of yellow-pigmented Gram-negative, rod-shaped, gamma-proteobacteria. The species in this genus have been isolated from many aquatic and terrestrial environments, in association with insects, animals and humans, as well as phytopathogens causing a variety of diseases in agricultural crop plants such as galls, wilting, soft rot and necrosis (1). *Pantoea stewartii* subsp. *stewartii* (*P. stewartii*, previously known as *Erwinia stewartii*) is the causal agent of Stewart's wilt in maize (2). This disease is a serious problem causing economic loss in corn production in the North America and impacting seed exports (2-4). In the USA, Stewart's wilt is known to occur in the midwest and mid-Atlantic states where susceptible corn varieties are cultivated (2, 5).

P. stewartii is transmitted to corn via an insect vector, the corn flea beetle *Chaetocnema pulicaria* (6). The bacteria can overwinter inside the gut of this insect and infect the corn seedlings during the feeding process of the beetles after they emerge in spring following winter hibernation in the soil (7). When beetle excrement gets deposited in wounds created by scratching, *P. stewartii* will enter the apoplast in the first phase of the disease leading to water-soaked lesions. In the second phase of the disease, the bacteria migrate to the xylem of the plant, grow to high cell density, and form a biofilm that blocks water transport, leading to wilt and death of the young seedlings. Several generations of beetles can be bred between the spring and fall, each acquiring the *P. stewartii* from infected plants and further spreading the bacteria across the fields. If more mature corn plants are infected, the bacterial disease is less severe and limited to the leaf blight symptoms of Stewart's wilt. However, these infected plants provide the source of bacteria for the infestation of the last generation of beetles prior to their winter hibernation. The bacterial transmission cycle will begin again once the hibernating beetles emerge the next

spring (2). Correa et al demonstrated that *P. stewartii* utilizes two type III secretion systems (T3SS) to colonize its two hosts, one for corn and the other for the beetle (7). These two distinct T3SS are located on two separate mega-plasmids among the multiple plasmids possessed by this phytopathogen (8).

P. stewartii is capable of quorum sensing (QS), a cell-to-cell communication phenomenon widely identified in eubacteria (9). *P. stewartii* QS is a function of the interaction between acyl-homoserine lactone (AHL), produced by a LuxI homologue named EsaI, and its cognate transcription factor EsaR, a homologue of LuxR, at high cell density. Unlike the canonical LuxI/R system, EsaR is a dual-level transcription factor that represses or activates its targets at low cell density in the absence of the interaction with AHL whereas derepression or deactivation occurs at high cell density when EsaR-AHL complexes form (10-14). It is established that QS plays an important role in the virulence of *P. stewartii* in corn since QS mutants are avirulent or have severely reduced virulence (12). This QS-mediated virulence is due to the regulation of *P. stewartii* exopolysaccharide (EPS) production (12) and surface motility (15). Two transcription factors directly controlled by EsaR, RcsA and LrhA, were found to play critical roles in the regulation of the EPS production and surface motility of *P. stewartii*, respectively (16-18). The other main pathogenicity factor of *P. stewartii* is controlled by the plant-colonization Hrp (**h**ypersensitive **r**esponse and **p**athogenicity) T3SS (2) that delivers protein effectors into the plant host cell upon direct contact. The expression of this system was found to be up-regulated during *in planta* growth compared to *in vitro* growth (19). In order to identify additional *P. stewartii* genes important for *in planta* colonization and growth, a Tn-Seq study was employed.

Tn-Seq is a powerful tool in which a library of transposon mutants is generated and grown under different test conditions prior to collecting DNA for high-throughput sequencing to identify the full set of genes necessary for viability under the test conditions (20, 21). The majority of Tn-Seq studies have most commonly utilized libraries generated by randomly inserted transposons from mariner-based transposons or Tn5-based vectors (21). This approach has been used to study essential genes under *in vitro* and *in vivo* conditions for a number of human and animal pathogens such as *Mycobacterium tuberculosis* (22), *Pseudomonas aeruginosa* (23, 24), *Streptococcus pyogenes* (25), and *Burkholderia pseudomallei* (26). However, the application of this approach in phytopathogens is very limited. One study used this approach to identify essential genes for *in vitro* growth of the endophyte *Herbaspirillum seropedicae* SmR1 that beneficially colonizes many crops, including rice, maize, sugarcane and sorghum (27). In addition, Cole et al. demonstrated the use of a randomly barcoded transposon mutagenesis sequencing method to establish a genome-wide map of bacterial genes required for colonization of the *Arabidopsis thaliana* root system by *Pseudomonas simiae* (28). Here, Tn-Seq has been applied to understand, at the genomic level, the genes that are critical for *in planta* colonization and survival of *P. stewartii* in corn.

MATERIALS AND METHODS

Strains and growth conditions

Strains used in this study are listed in Table 5.1. *P. stewartii* strains were grown in either Luria-Bertani (LB) (10 g/L tryptone, 5 g/L yeast extract, and 5 g/L NaCl) or M9 minimal medium (1X M9 salts, 2 mM $MgCl_2$, 0.1 mM $CaCl_2$ and 0.2% glucose) and *Escherichia coli* strains were grown in LB broth or on plates with 1.5% agar. The following antibiotics were supplemented in the growth medium as required: ampicillin (Ap, 100 μg/ml), chloramphenicol

(Cm, 35 μg/ml), kanamycin (Kn, 50 μg/ml), nalidixic acid (Nal, 30 μg/ml), or streptomycin (Str, 100 μg/ml). *E. coli* strains were grown at 37°C and *P. stewartii* strains were cultured at 30°C.

Estimation of the bottleneck effect in xylem infection assays

A critical issue with the Tn-Seq approach is the bottleneck effect during an infection assay (i.e. the number of bacteria that can effectively colonize the host in a single dose). Normally five μl of bacterial suspension (~3 x 10^5 colony forming units (CFU) of exponentially growing *P. stewartii* cell suspension at $OD_{600\ nm}$ = 0.2) is used to infect the corn seedlings in xylem infection virulence assays (18), also described below. Using a higher volume (e. g. 10 μl) of inoculum was problematic during the inoculation process. Therefore, instead of concentrating one ml of bacterial cell suspension at $OD_{600\ nm}$ = 0.2, ten ml were used. Three seedlings with a height of ~6-10 cm were inoculated with five μl (~3 x 10^6 CFU) for one hr and their stems were surface sterilized and sliced into ~1 mm pieces. Then, they were soaked in phosphate buffered saline (PBS, 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4 and 2 mM KH_2PO_4, pH 7.4) for 30 min and spread on LB Nal agar plates to enumerate the CFU. The experiment was done in duplicate and the bottleneck effect was estimated as the percentage of CFU_{output} per $CFU_{input.}$

Generation of the *P. stewartii* DC283 transposon mutant library

The *P. stewartii* DC283 transposon mutant library was generated via conjugation between an *E. coli* S17-1 donor strain containing plasmid pSAM-DKm (26) and the wild-type *P. stewartii* DC283 as a recipient strain. Vector pSAM-DKm contains a *mariner*-family transposon with a kanamycin resistance cassette and a *Himar1* C9 transposase with its upstream regulatory region. Overnight cultures of donor and recipient strains were subcultured in LB with the appropriate antibiotic supplemented to an $OD_{600\ nm}$ of 0.4-0.6 at 37ºC and 30ºC, respectively,

with shaking at 250 rpm. One ml of each strain per conjugation was harvested and washed twice with LB before mixing in 50 µl LB and then spotting on a LB agar plate. These conjugation plates were incubated at 30^{o}C overnight. Transposon mutants were selected on LB Nal Kn agar plates and grown for 48 hrs at 30^{o}C. Ten independent conjugations were conducted and colonies of mutants from ten plates were pooled to gain a library of ~4 x 10^4 mutants. The pool was aliquoted and mixed with glycerol prior to -70^{o}C storage. Independent transposon mutant pools from the library aliquots were prepared for use in the *in vitro* and *in planta* experiments performed in duplicate.

Tn-Seq screen for growth in LB medium

Each of the two pools of *P. stewartii* DC283 transposon mutants was grown overnight in 100 ml LB Kn at 30^{o}C with shaking. The cells were collected and washed twice with LB to inoculate 100 ml LB Kn medium to an $OD_{600\ nm}$ of 0.05. The mutant libraries were grown to an $OD_{600\ nm}$ of 0.2 (~6 x10^7 CFU/ml) at 30^{o}C with shaking at 250 rpm. Forty ml of bacterial suspension in LB were collected and centrifuged to harvest the cell pellet for DNA extraction to serve as a control for the Tn-Seq analysis, while ten ml of the suspension of mutant cells were grown in LB to an $OD_{600\ nm}$ of 0.2 and harvested for the plant infection. Experiments were done in duplicate using the two independently prepared transposon mutant pools.

Tn-Seq screen for growth in M9 minimal medium

The same overnight culture containing the mutant pool in LB, described above, was collected and washed twice with M9 medium to inoculate 100 ml M9 plus Kn medium to an $OD_{600\ nm}$ of 0.05. Bacterial mutants were allowed to grow in M9 Kn medium to an $OD_{600\ nm}$ of 0.2 at 30^{o}C with shaking at 250 rpm. Forty ml of bacterial suspension were collected and

centrifuged to harvest the cell pellet prior to DNA extraction for Tn-Seq analysis. Experiments were done in duplicate with the two independently prepared transposon mutant pools.

Tn-Seq screen for survival in infected plants

The pellet from the ten ml suspension of mutant cells growth in LB medium, described above, was washed twice with PBS and resuspended in one ml PBS prior to plant inoculation. Sweet corn seedlings (*Zea mays* cv. Jubilee, 2B Seeds) were grown on Sunshine mix #1 soil in a 30^{o}C growth chamber (Percival Scientific, Inc.) for seven days prior to inoculation with five µl of bacterial culture of mutants suspended in PBS based on modified published protocols (18). Thirty seedlings were surface cleaned with 70% ethanol prior to causing a ~1 cm wound using a needle to scratch the stem and deposit the bacterial suspension. The infected plants were grown to day six to retrieve the bacterial cells growing within them. The stems of day six post-infected plants were cut and surface sterilized with 70% ethanol before being cut into ~1 mm pieces that were immersed in PBS for 2 hrs to retrieve the bacteria from the xylem. The bacterial pellet was then subjected to DNA extraction for Tn-Seq analysis. Experiments were done in duplicate with the two independently prepared transposon mutant pools.

DNA extraction and Tn-Seq processing

Total DNA was extracted from duplicate samples of the three different growth conditions (i.e. LB, M9, and *in planta*) using a Qiagen DNeasy Blood & Tissue Kit per the manufacturer's recommendations for Gram-negative bacteria. DNA samples were then concentrated via ethanol precipitation. Four µg of DNA from each sample were used in a *Mme*I (NEB) digestion reaction per the manufacturer's instructions at 37^{o}C, overnight. The reactions were then purified with a QIAquick PCR purification kit (Qiagen) and eluted in 60 µl 10 mM Tris-HCl. A set of control

reactions with the same conditions as the *Mme*I digestion, but without addition of the enzyme were used as a control to check for nuclease contamination. Fifty ng of each digested DNA sample and control were loaded on a 1% agarose gel to visualize the results. The remaining samples of digested DNA were shipped to the Carver Biotechnology Center at the University of Illinois at Urbana-Champaign for processing and Hi-Seq Illumina sequencing.

Bioinformatics analysis

Raw Illumina HiSeq single 50 bp sequencing data were demultiplexed by the sequencing facility as six fastq files. Geneious R10.1 (Biomatters Ltd) was used to process and analyze these fastq files prior to enumerating the insertion numbers of the transposon mutant libraries. The raw reads from each fastq file were first mapped, using the 'Map to Reference' function in Geneious with no gap or mismatch allowed, to the specific sequence found on the mariner transposon in addition to the TA site at the 3' end, in bold, (GACCGGGGACTTATCATCCAACCTGT**TA**), where the insertion occurred. This subset of "Used Reads" was pulled out of the whole raw data files for further processing. Then, the first 27 bp from the 5' end of the reads that mapped to the transposon sequence above were trimmed to generate the reads that can be specifically mapped to the complete genome of *P. stewartii* DC283 (NCBI reference CP017581-92). The total reads resulted from the complete and incomplete *Mme*I digestion which normally cuts the DNA 18-20 bp from its recognition site (underlined in the above sequence). Only the short reads, ~16 bp, resulting from the trimming of the first 27 bp were used for the analysis in this study because the longer reads were either a consequence of the incomplete *Mme*I digestion or unspecific amplification from the used sequencing primer (ACACTCTTTCCCTACACGACGCTCTTCCGATCT). The ~16-bp reads were selected based on their size using function 'Extract' to select reads ≤ 16 bp in the folder containing the trimmed

reads. Then the trimmed and size-selected reads were mapped to the reference sequences (the main chromosome and 11 extra-chromosomal DNA sequences in *P. stewartii* DC283) using the 'Map to Reference' function with no gap and mismatch allowed and reads that could be matched in multiple locations were excluded by the selection of the "To None" option for repeated sequences. Finally, the "Calculate Expression Levels" command was used to determine raw read counts (RC) mapped to each coding sequence that were exported from Geneious for further analysis using Microsoft Excel.

Strain constructions for deletion and complementation strains

Primers used for strain construction are listed in Table 5.2. The two procedures used for strain generation have been previously described (18). Briefly, chromosomal deletions of *ompC*, *ompA* and *lon* were constructed by using Gateway technology (Life Technologies) to transfer a 2-kb DNA fragment, containing one kb from both up and down-stream of the gene of interest, from the cloning plasmids into the final suicide vector. Then, the suicide vectors were moved via conjugation into the wild-type *P. stewartii* DC283 strain. Double homologous recombination led to the generation of chromosomal deletion strains. In order to construct the chromosomal complementation strains of the corresponding deletion strains, coding sequences with their native upstream promoter regions were inserted into the neutral region downstream of *glmS* on the chromosome of the *P. stewartii* deletion strains with the pUC18R6K-mini-Tn7-cat vector (29) PCR products generated from constructed strains were sequenced to confirm the integrity of the constructs.

In planta competition assays

Plant growing conditions were the same as those used in the *in planta* Tn-Seq sample preparation described above. Corn seedlings at day seven post sowing were inoculated with five µl of a mixed population of a deletion strain (Nal^R) and its corresponding complementation strain (Nal^R/Cm^R), serving as the wild-type control, in a 1:1 ratio (or a 1:9 ratio). These strains were grown in the appropriate antibiotics in LB to an $OD_{600\ nm}$ of 0.2 prior to washing twice with PBS and combining in PBS solution with the appropriate volumes. Six plant seedlings were infected for each combination. After growing for six days, the stems of the infected plants were surface sterilized with ethanol 70% and sliced into small pieces to soak in PBS for total 2 hrs as described above for Tn-Seq sample processing. The bacterial cell pellet retrieved from each plant was serially diluted and spread on to LB agar with just Nal and Nal/Cm for colony enumeration of the deletion and complementation strains, respectively. Bacteria in each plant were enumerated in duplication and the average numbers from two replicates were used to calculate the relative competition index (RCI = ([CFU of mutant/CFU of complementation]$_{output}$)/([CFU of mutant/CFU of complementation]$_{input}$)). For experiments with a 1:9 ratio, 100 colonies grown in LB Nal were patched on LB Nal/Cm first and then on LB Nal to evaluate the RCI.

Plant xylem virulence assays

For virulence assays (18), sweet corn seeds were sowed in Promix soil for six days with the same growing conditions described above. Each tested strain was grown to an $OD_{600\ nm}$ of 0.2 in LB prior to washing with PBS. Five µl of cell suspension were inoculated into an ~1 cm wound, caused by a needle incision at ~1 cm above the soil line, on the stem of each plant. Fifteen plants were used for each treatment to measure the average disease severity score (0 = no symptoms; 1 = few scattered lesions; 2 = scattered water soaking symptoms; 3 = numerous

lesions and slight wilting; 4 = moderately severe wilt; 5 = death) at day 12 post-infection. The wild-type and PBS controls were included in each experiment and the accumulative numbers of these plants were used for the comparison. A Student's T-test with $p < 0.05$ between tested strain and wild-type treatment was considered statistically significant.

Data availability

The raw data generated from this study will be deposited in the NCBI Sequence Read Archive (SRA).

RESULTS

Generation of the *P. stewartii* DC283 transposon mutant library for high-throughput Tn-sequencing analysis

The overall study design is depicted in Figure 5.1. A transposon insertion library of ~40,000 mutants was generated to identify genes contributing to the survival of *P. stewartii* under three growth conditions, two *in vitro* and one *in planta*. Two independent pools of the library were separately grown in LB medium overnight to be revived from frozen aliquots. These two pools were then subcultured in two different nutrient sources: LB, a rich undefined medium, and M9 minimal medium, containing a defined amount of several inorganic salts and a low percentage of glucose. The LB grown library pool was used to inoculate 30 corn seedlings to determine genes required for *in planta* survival of *P. stewartii*.

The raw sequencing data were demultiplexed by the sequencing facility prior to analysis of the high-quality sequencing data. The total numbers of sequencing reads for each library pool were similar. Pool 1 and pool 2 from LB growth contained 21,626,331 and 23,320,367 reads, respectively. The two pools from growth in M9 minimal medium had 21,212,131 and 24,689,688

reads. Sequencing data from the two *in planta* growth pools resulted in 26,977,393 and 20,833,198 reads, respectively. The original sequencing reads with 50-bp length were first mapped to a specific sequence found on the transposon plus the TA site, (GACCGGGGACTTATCATCCAACCTGT**TA**), to select for the most likely inserted sequences for further analysis. Once the transposon insert reads were selected, they were then truncated to contain just the ~16 bp *P. stewartii* sequence specific to the insertion sites. Only then were these reads mapped to the complete annotated genome of *P. stewartii* DC283 (8) to be enumerated for each coding sequence (CDS) on the main chromosome (NCBI reference CP017581). There were 1,235,350 and 1,766,200 read counts (RC) that mapped to the coding sequences in CP017581 for the two pools of library grown in LB at an $OD_{600\ nm}$ of 0.2, respectively. The replicates grown in M9 medium resulted in 1,291,974 and 1,702,229 RC, respectively. The correlation of the transposon insertion counts for the two replicates is very high for samples grown in LB and in M9 (Fig. 5.2A & 5.2B).

However, the RC to coding regions for the library growth in plant were only 499,445 (40.4% of LB pre-inoculum sample) for pool 1 and 342,995 (19.4% of LB pre-inoculum sample) for pool 2. Moreover, there is a lower correlation between the two replicates of *in planta* samples (Fig. 5.2C) versus the LB and M9 samples. Both of these effects likely reflect the role of the infection bottleneck in the plants. Because of the large differences in the *in planta* results, the RC for each coding region from the two replicates were not averaged but instead were kept separate for the independent, pair-wise comparison for each transposon mutant library data set.

The whole genome assembly of *P. stewartii* DC283 (8) showed that this bacterium contains multiple plasmids with many copies except the largest plasmid, pDSJ010 (NCBI reference CP017591) with 304,641 bp, that appears to have a single copy in the cell. Therefore,

Tn-Seq analysis was also utilized to predict the critical genes on this plasmid using the same set of criteria applied for the chromosome. Overall, there were 92,696 and 135,918 RC in LB samples, 122,905 and 165,380 RC in M9 samples, and 46,602 and 58,436 RC *in planta* growth samples, respectively corresponding to pDSJ010.

Identification of putative *P. stewartii* genes essential for growth in LB medium by Tn-Seq

To determine candidate essential genes, those genes with the lowest RC were identified from the LB pools. In principle an absolutely essential gene should not tolerate any transposon insertions and would have an RC of 0. However, genes with duplicated sequence could appear to exhibit transposon inserts if only one copy is essential. Therefore, to simplify the analysis, all of the ~16 bp reads which mapped to multiple locations in the genome were excluded from the counts. There were 168 genes in the bacterial chromosome with 0 RC identified in both LB pools. Of the 4399 annotated genes in the chromosome, these genes are thought to be absolutely essential for the viability of *P. stewartii* in LB medium (Table C.1). Despite the close correlation of the overall Tn-Seq data between the two LB samples, there are large differences in the 397 genes with $0 < RC < 10$ in these individual LB grown transposon insertion pools (Figure C.2). There was a total of 565 genes from the chromosome of *P. stewartii* DC283 with less than 10 RC in both LB grown pools (Table C.1). These genes are also potentially important for bacterial survival in LB. In addition, there were nine genes with less than 10 RC found amongst the 297 annotated genes in pDSJ010 (Table C.2), but no genes with zero RC in both samples.

Identification of genes with mutations differentially present in M9 minimal medium compared to LB medium

All coding sequences with greater than or equal to 10 RC in both LB samples were compared to the RC of the same coding sequenes in the mutant library grown in M9 minimal medium. Using a two-fold cut-off for both pair-wise comparisons, a list of 77 genes with higher insertion mutants in LB than in M9 medium was generated (Table C.3). The majority of these genes are involved in the purine biosynthesis pathway (*pur* genes) or synthesis of amino acids (e.g. *cys*, *leu*, *arg* genes). These genes are probably important for growth in the minimal medium where the external source of these important metabolites is absent. Interestingly, there was one gene (*rbsR* encoding a ribose transport operon repressor) found to have better fitness with its mutant version in the M9 minimal medium. The RC for this gene were about 2.1-fold higher in M9 compared to growth in LB (Table C.3). This suggests that the membrane transport system for ribose is required for growth in M9 medium since the mutant of its repressor had a growth benefit in this medium. Separately, a single gene (DSJ_25325), encoding a hypothetical protein, on the plasmid pDSJ010 was found to have two-fold better fitness in M9 medium when its gene had a transposon insertion in it (Table C.3).

Identification of *P. stewartii* potential crucial genes for survival in infected plants

The putatively non-essential genes for growth in LB, with more than 10 RC in both LB samples, were compared with the RC in the pools of DNA from 30 infected seedlings. The total number of coding sequences with greater than 10 RC in both LB duplicates was divided into two lists: (i) the genes with RC in both LB duplicates greater than 100, and (ii) the genes with $10 < RC < 100$. Due to the abundance of the starting mutants in LB samples for the first list with higher RC, this provided higher confidence in predicting the role of these genes during *in planta*

colonization and survival compared to the second list. Genes with more than ten-fold overall difference in the RC between LB samples and *in planta* samples were considered to be significant for the analysis since the estimated bottleneck effect is ~10% (Figure C.1). Any genes with important roles during growth in M9 minimal medium (Table C.3) were excluded from the list of genes uniquely important for *in planta* growth.

It was estimated that the bacteria grew for approximately ten generations in both the LB and *in planta* growth conditions (data not shown). Genes with lower RC *in planta* growth than in LB medium were considered to be important for *in planta* colonization and survival since their mutants cannot compete well inside the plants. There were 193 genes predicted to belong to this category (Table C.4) with the RC in LB samples greater than 100, among them 47 genes had less than 10 RC in plant samples (Table 5.3). There were 184 genes with the RC in LB samples between 10 to 100 RC with greater than ten-fold reduction of mutants *in planta* samples compared to in LB growth (Table C.5), among them 26 genes had zero RC in plant samples (Table 5.4). Moreover, ten genes with RC greater than 10 in LB samples on the plasmid pDSJ010 were recognized to have ten-fold reduction in plant samples (Table 5.5).

Genes with higher RC *in planta* growth compared to LB samples were considered to be restricted in the wild-type *P. stewarti* for *in planta* colonization and survival since the mutants of these genes provided better fitness for the bacterial *in planta* growth. Seven genes in the *P. stewartii* chromosome were found in this group; only one of them has a RC in the LB samples lower than 100 (*rcsA*) (Table 5.6). No genes on the plasmid pDSJ010 were identified to have higher RC *in planta* growth up to ten-fold.

Validation of Tn-Seq data using *in planta* competition assays

LrhA and RcsA have previously been shown to be important factors for *in planta* virulence in *P. stewartii* (18). Their deletion mutants and the corresponding complementation strains were available to conduct initial competition assays for confirmation of Tn-Seq analysis. Transposon insertion mutants of *lrhA* showed very low RC in the pools of DNA samples retrieved from *in planta* infection (Table C.5) whereas *rcsA* displayed a higher number of insertion mutants *in planta* growth compared to the growth in LB medium (Table 5.6). In addition, three more genes from the top list of genes important for *in planta* growth (Table 5.3) were chosen to generate the deletion and the complementation strains to further validate the Tn-Seq screening. Each of these genes produces a monocistronic transcript. The amounts of these transcripts were checked to ensure the genes are expressed *in planta* using previously published RNA-Seq data (19). Consequently, *ompC* (DSJ_09635, encoding a general porin OmpC or OmpF), *ompA* (DSJ_09745, encoding a porin OmpA), and *lon* (DSJ_19160, encoding an endopeptidase) were selected for validation.

The deletion and complementation strains for five genes (i.e. *rcsA*, *lrhA*, *ompC*, *ompA*, and *lon*) were tested in a direct competition at a 1:1 ratio between the deletion strain and the complementation strain, representative of the wild-type strain, to test predictions from the Tn-Seq analysis. The relative competition indices from these assays were shown in Figure 5.3. The competition assay results showed that the *rcsA* mutant has higher fitness whereas the *ompC*, and *lon* mutants have lower fitness as initially shown by the Tn-Seq data. However, *lrhA* and *ompA* have fitness levels close to the wild-type strain in these assays. The assays for *ompA* were repeated with the same 1:1 ratio and a lower concentration of the *ompA* deletion strain in a 1:9

ratio to reexamine the competition assays for this gene. Relative competition indices from these repeated assays were similar to the results from the first trial (Fig. C.3).

Evaluate the contribution of *ompC, ompA* and *lon* to the *in planta* virulence of *P. stewartii*

Assays were conducted to assess the contribution of *ompC, ompA*, and *lon* to the virulence of *P. stewartii* using the xylem infection model (Fig. 5.4). Similar assays has previously been performed with *lrhA* and *rcsA* (18, 54). The deletion strain of *ompC* (Δ*ompC*) had a lower score of disease severity compared to the wild-type infected plants ($p < 0.057$ in a pair-wise comparison with the wild-type treatment using a Student's T-test) while its complementation strain (Δ*ompC/ompC*$^+$) restored the phenotype close to the wild-type level. Similarly, plants infected with Δ*ompA* displayed statistically significant disease severity reduction ($p < 0.05$) whereas its complementation strain (Δ*ompA/ompA*$^+$) partially recovered the wild-type disease score. The disease severity score of plants infected with Δ*lon* were also statistically significantly reduced compared to the average score of the wild-type treatment ($p < 0.05$). However, the score of plants infected with the complementation strain (Δ*lon/lon*$^+$) was not reestablished to the level of the wild-type infected plants. As a check, a separate phenotypic examination of the colony morphology of the wild-type, Δ*lon*, and Δ*lon/lon*$^+$ strains appears to confirm the successful complementation of the deletion strain (Fig. C.4). In this assay, Δ*lon* showed a larger colony morphology whereas the wild-type strain and the complementation strain had a smaller size, similar to one another.

DISCUSSION

A Tn-Seq transposon library was generated to investigate *P. stewartii* genes important for growth under different conditions, with a particular emphasis on establishing the genes critical

for *in planta* colonization and survival. A reduced number of transposon insertions or RC in a gene recovered from the *in planta* samples (e.g. *ompC*) indicates the gene is important in growth/survival, whereas a higher number of transposon insertion or RC (e.g. *rcsA*) indicates that the wild-type gene actually reduces the bacterial fitness since a mutant form of it provides enhanced fitness (Fig. C.5). The size of the transposon library used for the studies was moderate to insure identification of genes important for *in vitro* growth with a minimized bottleneck effect during the *in planta* infection. There are 168 genes with no transposon insertions found in the library recovered from growth in LB suggesting their essential role for growth in this rich medium. Among these, the products of 71 genes are annotated as hypothetical proteins suggesting the importance of the undetermined functions encoded by these genes in survival of the bacterium. An additional 397 genes having less than 10 putative insertions in the LB growth condition might also play an important role in the survival of *P. stewartii*. The annotation of these genes showed that they belong to many fundamental processes including replication, transcription, translation, protein secretion, cell division, peptidoglycan synthesis, and amino acids synthesis. The transposon library grown in LB was considered the control library to identify genes important for other growth conditions. Therefore, the 565 genes with less than 10 transposon insertions grown in LB were excluded from comparisons with the M9 and *in planta* samples. Of the remaining 3,834 *P. stewartii* genes dispensable for growth in LB, only 77 genes were found to be two-fold reduced in the number of mutants grown in M9 minimal medium. This finding suggests that *P. stewartii* does not require a large repertoire of genes to grow in this minimal medium or that the cells were not grown for a sufficient number of generations to see larger physiological impacts. These genes were also excluded from the analysis of genes important for growth inside the plant host because strains with mutations in them might be less

fit in this environment due to the inability to synthesize the required nutrients rather than the direct effect caused by the host response.

Overall, 377 genes were found to have more than ten-fold less transposon insertions following *in planta* growth than following growth in LB. This indicates that these genes are likely important for survival inside the plant host. The majority of these genes code for outer membrane proteins, transporters, metabolic enzymes, iron uptake and metabolism, transcriptional regulators, capsule synthesis, and hypothetical proteins. Three gene deletion strains (i.e. Δ*ompC*, Δ*ompA*, and Δ*lon*) were constructed in this study. In addition, a previously constructed strain of interest with a deletion in *lrhA* was used to validate the findings from the Tn-Seq analysis using a conventional competition assay where the complementation strains served as the wild-type control. The overall trend of the Tn-Seq prediction for *ompC* and *lon* were confirmed by 1:1 ratio competition assays. The competition assay results for *lrhA* and *ompA* gene were different from the Tn-Seq prediction for the inability to compete *in planta* for *lrhA* and *ompA* transposon insertion mutants. With regard to *lrhA*, the Tn-Seq results were not as compelling as for the other three genes. Thus, it was not predicted to play as strong role in *in planta* growth. The results with *ompA* were more surprising. However, the 1:1 inoculation ratio used in the competition assays does not truly reflect the ratio of these mutants in the actual Tn-Seq library used to inoculate the plants. Therefore, a 1:9 ratio was also employed for the *ompA* deletion strain, but this, too, did not demonstrate a reduction in the gene fitness *in planta*. An even greater difference in inoculation levels between the mutant and complement may be necessary to observe *in planta* differential survival, especially if there were shared products involved. It is also possible that the *ompA* transposon insertion mutants may have imposed undesirable polar effects that the markerless deletion mutant strain did not have, although it was

predicted to produce a monocistronic transcript. Nevertheless, the *in planta* virulence assays for *ompC*, *ompA,* and *lon* mutants and their complementation strains indicated that they each partially contribute to the virulence of *P. stewartii*. OmpC and Lon may contribute to the bacterial virulence due to their involvement in bacterial colonization and growth; whereas, OmpA is likely involved in virulence of *P. stewartii* by a different, as of yet un-identified mechanism.

Bacterial outer membrane proteins (OMPs) form channels in the outermost layer of the Gram-negative bacterial cell wall to serve as the permeability gates for nutrient uptake and waste excretion. OMPs also function as surface receptors, contribute to cell adherence to other surfaces, and maintain the structural integrity of the cell (30, 31). Approximately 3% of the Gram-negative bacterial genome is estimated to encode OMPs (32), and they are some of the most highly expressed proteins in the cell (31). Only two OMPs, LptD that is responsible for the insertion of lipopolysaccharide (LPS) into the outer membrane (OM) and BamA that is a component of the β-barrel assembly machinery (BAM) complex, were found to be essential in *Escherichia coli* (33, 34). Although these genes were not considered to be essential in LB growth in *P. stewartii*, other genes in the process of LPS insertion into the OM (LptA, LptC, LptF, LptG, LptE) were recognized as being essential by this Tn-Seq analysis (Table C.1).

The OM in Gram-negative bacteria is the first line of defense against toxic compounds since this membrane is impermeable to large, charged molecules. OMPs, called general porins, are aqueous open gates that allow the passive penetration of small hydrophobic molecules through diffusion across this barrier (35, 36). In *E. coli*, OmpC and OmpF are two general porins with ~ 10^5 copies per cell that function as general gateways for the entry of < 600 Da small molecules (37). Mutations in *ompC* are known to be involved in antibiotic resistance in several

Gram-negative pathogens, such as *E. coli* (38), *Klebsiella pneumonia* (39), and *Enterobacter aerogenes* (40). *E. coli* OmpC was also shown to be a target of the innate immunity response during infection since it is a lactoferrin binding protein (41) and anti-OmpC antibody was detected in human sera (38). However, the physical role of OmpC in *P. stewartii* is largely undefined. This work has demonstrated that OmpC plays a previous unappreciated role in the virulence of this phytopathogen.

Similarly, the function of OmpA in *P. stewartii* is also understudied. However, OmpA is one of the most well-studied OMPs in *E. coli* with both structural and ion-permeable porin functions (42). *E. coli* strains with *ompA* mutations showed lower survival rates in stress conditions such as high osmolarity, an acidic environment, and pooled human serum, but survived better inside cultured brain microvascular endothelial cells (43). OmpA is also involved in phenol resistance in *E. coli* (44) and its expression increases in biofilms (45). *K. pneumonia* OmpA can act as a microbial-associated molecular pattern (MAMP) that interacts with some immune cells like macrophages, natural killer cells and dendritic cells (46). In addition, OmpA family proteins have major roles in pathogenesis such as bacterial adhesion, cellular invasion, intracellular survival, and evasion of host defenses in a number of human pathogens, including *Acinetobacter baumannii*, *E. coli* meningitic strains, *K. pneumonia*, and *Neisseria gonorrhoeae* (47). It is unclear why the competition assay for *ompA* did not support the initial findings from Tn-Seq analysis in *P. stewartii*. Perhaps, it was a difference in the genetic nature of the two constructs or the percentage of the mutation in a mixed population. However, this study demonstrated the involvement of OmpA in the virulence of this phytopathogen. Future work will be necessary to determine the molecular and physiological basis of this finding.

The endopeptidase La, commonly known as Lon, is a highly conserved ATP-dependent protease in archaea and bacteria as well as eukaryotic mitochondria and peroxisomes. Lon plays a critical role in cellular homeostasis by mediating the degradation of damaged or aberrant polypeptides and short-lived metabolic regulation proteins (48). RcsA, a short-lived metabolic regulatory activator of capsule production, is a target of Lon degradation in *E. coli* and *P. stewartii* (49, 50). The Lon protease was also found to be essential for full virulence of *Pseudomonas aeruginosa* in multiple infection models (51). Experiments in this study demonstrated that *P. aeruginosa lon* mutants had impacts on expression of T3SS genes, motility and biofilm formation which might be reflected by the *in vivo* virulence reduction. In a systemic *Salmonella enterica* serovar Typhimurium infection of mice, Lon was found to be critically important for intracellular survival in murine macrophages and proliferation within the spleen causing systemic disease (52). The inability of *lon* mutant to compete during both the Tn-Seq analysis and the competition assays in the current study suggests that Lon is involved in the virulence of *P. stewartii* via the contribution to the *in planta* growth and survival, in addition to its role in the regulation of RcsA.

Intriguingly, mutants in a set of seven genes were identified to proliferate better *in planta* during the Tn-Seq investigation. Four of these genes code for regulators of capsule production (i.e. RcsABCD) in *P. stewartii* and three of them encode hypothetical proteins. Competition assays using a *rcsA* mutant confirmed that the strain was more fit *in planta* as had been predicted by Tn-Seq data. Capsule production is energetically expensive and is a common good that may be shared between neighboring cells. Thus, cheater cells that rely on the capsule produced by others will have a competitive advantage. On the other hand, many genes involved in the capsule biosynthetic pathway were found by Tn-Seq analysis as contributing factors of *in planta* growth

and survival (Tables 5.3, 5.4, C.5, and C.6) suggesting that capsule production itself is an indispensable component of bacterial successful proliferation *in planta*. The role of EPS/stewartan in virulence has previously been well established (2, 53). Its regulation is tightly controlled in *P. stewartii.* Any increase in levels, such as in a QS mutant where EsaR is inactive (12) or a *lon* mutant as shown here, or decrease in levels, such as a *rcsA* mutant (54) results in reduced virulence of *P. stewartii*. Thus, the process of capsule synthesis is a costly metabolism requiring a precise regulation for successful colonization and virulence of the phytopathogen.

In conclusion, Tn-Seq is a powerful tool to study genes important for select environmental conditions. However, Tn-Seq is only capable of identifying mutations with cell-autonomous phenotypes, which mean that the phenotypes caused by these mutations cannot be hidden by other cells in the population. For instance, QS mutants were not identified by the Tn-Seq analysis probably because the diffusible QS signal is publicly available and can be produced by other cells in the population therefore masking the phenotypes of a QS mutant. Despite this limitation, Tn-Seq has been successfully used to identify a set of genes important for the *in planta* survival of the wilt-pathogen *P. stewartii.* The genes *ompC* and *lon* are needed for *P. stewartii* to colonize and grow *in planta*, while the loss of *rcsA* increases the fitness of the bacterium. Lon was also shown to play a role in virulence, perhaps due to the *in planta* growth defect it exhibits. OmpA was discovered to play a role in virulence that needs to be further explored. Future studies to define the role of other genes identified through the Tn-Seq approach as playing a function in *in planta* survival will promote a greater understanding of host-microbe interactions among different phytopathogens, especially xylem-dwelling bacteria causing wilt disease.

ACKNOWLEDGMENTS

This work was funded by the Department of Biological Sciences and Virginia Tech Life Sciences 1 Building Fund (AMS), the Fralin Institute Genomics Analysis Incubator (RVJ), and a GSDA Fellowship and GRDP Award from the Graduate School (DAD). We are grateful to D. Yoder-Himes and J. Warawa for sharing plasmids and providing advice about generating the mutant library. We thank the laboratories of R. Helm and B. Winkel for sharing their plant growth chamber to conduct all the *in planta* assays, as well as H. Packard, B. Thomas, and S. Williams for assistance with competition assays, and H. Packard for assistance with the plant virulence assays.

REFERENCES

1. **Walterson AM, Stavrinides J.** 2015. *Pantoea:* insights into a highly versatile and diverse genus within the *Enterobacteriaceae*. FEMS Microbiology Reviews **39:**968-984.
2. **Roper MC.** 2011. *Pantoea stewartii* subsp. *stewartii*: lessons learned from a xylem-dwelling pathogen of sweet corn. Molecular Plant Pathology **12:**628-637.
3. **Tambong JT.** 2015. Specific identification and detection of *Pantoea stewartii* subsp. *stewartii* using a membrane-based multi-gene oligonucleotide array. Canadian Journal of Plant Pathology **37** 414-426.
4. **Pataky J, Ikin R.** 2003. Pest risk analysis: the risk of introducing *Erwinia stewartii* in maize seed. The International Seed Federation, Nyon, Switzerland.
5. **Freeman ND, Pataky JK.** 2001. Levels of Stewart's wilt resistance necessary to prevent reductions in yield of sweet corn hybrids. Plant Disease **85:**1278–1284.
6. **Esker PD, Nutter FW.** 2002. Assessing the risk of Stewart's disease of corn through improved knowledge of the role of the corn flea beetle vector. Phytopathology **92:**668-670.
7. **Correa VR, Majerczak DR, Ammar el D, Merighi M, Pratt RC, Hogenhout SA, Coplin DL, Redinbaugh MG.** 2012. The bacterium *Pantoea stewartii* uses two different type III secretion systems to colonize its plant host and insect vector. Applied and Environmental Microbiology **78:**6327-6336.
8. **Duong DA, Stevens AM, Jensen RV.** 2017. Complete genome assembly of *Pantoea stewartii* subsp. *stewartii* DC283, a corn pathogen. Genome Announcements **5**: e00435-17.
9. **Rutherford ST, Bassler BL.** 2012. Bacterial quorum sensing: its role in virulence and possibilities for its control. Cold Spring Harbor Perspectives in Medicine **2**: a012427.
10. **Minogue TD, Wehland-von Trebra M, Bernhard F, von Bodman SB.** 2002. The autoregulatory role of EsaR, a quorum-sensing regulator in *Pantoea stewartii* ssp. *stewartii*: evidence for a repressor function. Molecular Microbiology **44:**1625-1635.
11. **Beck von Bodman S, Farrand SK.** 1995. Capsular polysaccharide biosynthesis and pathogenicity in *Erwinia stewartii* require induction by an N-acylhomoserine lactone autoinducer. Journal of Bacteriology **177:**5000-5008.
12. **von Bodman SB, Majerczak DR, Coplin DL.** 1998. A negative regulator mediates quorum-sensing control of exopolysaccharide production in *Pantoea stewartii* subsp. *stewartii*. Proceedings of the National Academy of Sciences of the United States of America **95:**7687-7692.
13. **von Bodman SB, Ball JK, Faini MA, Herrera CM, Minogue TD, Urbanowski ML, Stevens AM.** 2003. The quorum sensing negative regulators EsaR and $ExpR_{Ecc}$, homologues within the LuxR family, retain the ability to function as activators of transcription. Journal of Bacteriology **185:**7001-7007.
14. **Schu DJ, Ramachandran R, Geissinger JS, Stevens AM.** 2011. Probing the impact of ligand binding on the acyl-homoserine lactone-hindered transcription factor EsaR of *Pantoea stewartii* subsp. *stewartii*. Journal of Bacteriology **193:**6315-6322.
15. **Herrera CM, Koutsoudis MD, Wang X, von Bodman SB.** 2008. *Pantoea stewartii* subsp. *stewartii* exhibits surface motility, which is a critical aspect of Stewart's wilt disease development on maize. Molecular Plant-Microbe Interactions **21:**1359-1370.
16. **Minogue TD, Carlier AL, Koutsoudis MD, von Bodman SB.** 2005. The cell density-dependent expression of stewartan exopolysaccharide in *Pantoea stewartii* ssp. *stewartii* is a function of EsaR-mediated repression of the *rcsA* gene. Molecular Microbiology **56:**189-203.

17. **Ramachandran R, Stevens AM.** 2013. Proteomic analysis of the quorum-sensing regulon in *Pantoea stewartii* and identification of direct targets of EsaR. Applied and Environmental Microbiology **79:**6244-6252.
18. **Kernell Burke A, Duong DA, Jensen RV, Stevens AM.** 2015. Analyzing the transcriptomes of two quorum-sensing controlled transcription factors, RcsA and LrhA, important for *Pantoea stewartii* virulence. PLOS ONE **10:**e0145358.
19. **Packard H, Kernell Burke A, Jensen RV, Stevens AM.** 2017. Analysis of the *in planta* transcriptome expressed by the corn pathogen *Pantoea stewartii* subsp. *stewartii* via RNA-Seq. PeerJ **5:**e3237.
20. **van Opijnen T, Bodi KL, Camilli A.** 2009. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. Nature Methods **6:**767-772.
21. **Chao MC, Abel S, Davis BM, Waldor MK.** 2016. The design and analysis of transposon insertion sequencing experiments. Nature Reviews. Microbiology **14:**119-128.
22. **Zhang YJ, Ioerger TR, Huttenhower C, Long JE, Sassetti CM, Sacchettini JC, Rubin EJ.** 2012. Global assessment of genomic regions required for growth in *Mycobacterium tuberculosis*. PLOS Pathogens **8:**e1002946.
23. **Skurnik D, Roux D, Aschard H, Cattoir V, Yoder-Himes D, Lory S, Pier GB.** 2013. A comprehensive analysis of *in vitro* and *in vivo* genetic fitness of *Pseudomonas aeruginosa* using high-throughput sequencing of transposon libraries. PLOS Pathogens **9:**e1003582.
24. **Turner KH, Wessel AK, Palmer GC, Murray JL, Whiteley M.** 2015. Essential genome of *Pseudomonas aeruginosa* in cystic fibrosis sputum. Proceedings of the National Academy of Sciences of the United States of America **112:**4110-4115.
25. **Le Breton Y, Belew AT, Valdes KM, Islam E, Curry P, Tettelin H, Shirtliff ME, El-Sayed NM, McIver KS.** 2015. Essential Genes in the Core Genome of the Human Pathogen *Streptococcus pyogenes*. Scientific Reports **5:**9838.
26. **Gutierrez MG, Yoder-Himes DR, Warawa JM.** 2015. Comprehensive identification of virulence factors required for respiratory melioidosis using Tn-seq mutagenesis. Frontiers in Cellular and Infection Microbiology **5:**78.
27. **Rosconi F, de Vries SP, Baig A, Fabiano E, Grant AJ.** 2016. Essential genes for *in vitro* growth of the endophyte *Herbaspirillum seropedicae* SmR1 as revealed by transposon insertion site sequencing. Applied and Environmental Microbiology **82:**6664-6671.
28. **Cole BJ, Feltcher ME, Waters RJ, Wetmore KM, Mucyn TS, Ryan EM, Wang G, Ul-Hasan S, McDonald M, Yoshikuni Y, Malmstrom RR, Deutschbauer AM, Dangl JL, Visel A.** 2017. Genome-wide identification of bacterial plant colonization genes. PLOS Biology **15:**e2002860.
29. **Choi KH, Gaynor JB, White KG, Lopez C, Bosio CM, Karkhoff-Schweizer RR, Schweizer HP.** 2005. A Tn7-based broad-range bacterial cloning and expression system. Nature Methods **2:**443-448.
30. **Rollauer SE, Sooreshjani MA, Noinaj N, Buchanan SK.** 2015. Outer membrane protein biogenesis in Gram-negative bacteria. Philosophical transactions of the Royal Society of London. Series B, Biological Sciences **370**.
31. **DiRienzo JM, Nakamura K, Inouye M.** 1978. The outer membrane proteins of Gram-negative bacteria: biosynthesis, assembly, and functions. Annual Review of Biochemistry **47:**481-532.
32. **Wimley WC.** 2003. The versatile beta-barrel membrane protein. Current Opinion in Structural Biology **13:**404-411.
33. **Wu T, McCandlish AC, Gronenberg LS, Chng SS, Silhavy TJ, Kahne D.** 2006. Identification of a protein complex that assembles lipopolysaccharide in the outer membrane of *Escherichia coli*. Proceedings of the National Academy of Sciences of the United States of America **103:**11754-11759.

34. **Albrecht R, Schutz M, Oberhettinger P, Faulstich M, Bermejo I, Rudel T, Diederichs K, Zeth K.** 2014. Structure of BamA, an essential factor in outer membrane protein biogenesis. Acta Crystallographica. Section D, Biological Crystallography **70:**1779-1789.
35. **Masi M, Pages JM.** 2013. Structure, Function and Regulation of Outer Membrane Proteins Involved in Drug Transport in Enterobactericeae: the OmpF/C - TolC Case. The Open Microbiology Journal **7:**22-33.
36. **Galdiero S, Falanga A, Cantisani M, Tarallo R, Della Pepa ME, D'Oriano V, Galdiero M.** 2012. Microbe-host interactions: structure and role of Gram-negative bacterial porins. Current Protein & Peptide Science **13:**843-854.
37. **Molloy MP, Herbert BR, Slade MB, Rabilloud T, Nouwens AS, Williams KL, Gooley AA.** 2000. Proteomic analysis of the *Escherichia coli* outer membrane. European Journal of Biochemistry **267:**2871-2881.
38. **Liu YF, Yan JJ, Lei HY, Teng CH, Wang MC, Tseng CC, Wu JJ.** 2012. Loss of outer membrane protein C in *Escherichia coli* contributes to both antibiotic resistance and escaping antibody-dependent bactericidal activity. Infection and Immunity **80:**1815-1822.
39. **Garcia-Fernandez A, Miriagou V, Papagiannitsis CC, Giordano A, Venditti M, Mancini C, Carattoli A.** 2010. An ertapenem-resistant extended-spectrum-beta-lactamase-producing *Klebsiella pneumoniae* clone carries a novel OmpK36 porin variant. Antimicrobial Agents and Chemotherapy **54:**4178-4184.
40. **Thiolas A, Bornet C, Davin-Regli A, Pages JM, Bollet C.** 2004. Resistance to imipenem, cefepime, and cefpirome associated with mutation in Omp36 osmoporin of *Enterobacter aerogenes*. Biochemical and Biophysical Research Communications **317:**851-856.
41. **Sallmann FR, Baveye-Descamps S, Pattus F, Salmon V, Branza N, Spik G, Legrand D.** 1999. Porins OmpC and PhoE of *Escherichia coli* as specific cell-surface targets of human lactoferrin. Binding characteristics and biological effects. The Journal of Biological Chemistry **274:**16107-16114.
42. **Hong H, Szabo G, Tamm LK.** 2006. Electrostatic couplings in OmpA ion-channel gating suggest a mechanism for pore opening. Nature Chemical Biology **2:**627-635.
43. **Wang Y.** 2002. The function of OmpA in *Escherichia coli*. Biochemical and Biophysical Research Communications **292:**396-401.
44. **Zhang DF, Li H, Lin XM, Wang SY, Peng XX.** 2011. Characterization of outer membrane proteins of *Escherichia coli* in response to phenol stress. Current Microbiology **62:**777-783.
45. **Orme R, Douglas CW, Rimmer S, Webb M.** 2006. Proteomic analysis of *Escherichia coli* biofilms reveals the overexpression of the outer membrane protein OmpA. Proteomics **6:**4269-4277.
46. **Chalifour A, Jeannin P, Gauchat JF, Blaecke A, Malissard M, N'Guyen T, Thieblemont N, Delneste Y.** 2004. Direct bacterial protein PAMP recognition by human NK cells involves TLRs and triggers alpha-defensin production. Blood **104:**1778-1783.
47. **Confer AW, Ayalew S.** 2013. The OmpA family of proteins: roles in bacterial pathogenesis and immunity. Veterinary Microbiology **163:**207-222.
48. **Lee I, Suzuki CK.** 2008. Functional mechanics of the ATP-dependent Lon protease- lessons from endogenous protein and synthetic peptide substrates. Biochimica et Biophysica acta **1784:**727-735.
49. **Torres-Cabassa AS, Gottesman S.** 1987. Capsule synthesis in *Escherichia coli* K-12 is regulated by proteolysis. Journal of Bacteriology **169:**981-989.
50. **Torres-Cabassa A, Gottesman S, Frederick RD, Dolph PJ, Coplin DL.** 1987. Control of extracellular polysaccharide synthesis in *Erwinia stewartii* and *Escherichia coli* K-12: a common regulatory function. Journal of Bacteriology **169:**4525-4531.

51. **Breidenstein EB, Janot L, Strehmel J, Fernandez L, Taylor PK, Kukavica-Ibrulj I, Gellatly SL, Levesque RC, Overhage J, Hancock RE.** 2012. The Lon protease is essential for full virulence in *Pseudomonas aeruginosa*. PLOS ONE **7:**e49123.
52. **Takaya A, Suzuki M, Matsui H, Tomoyasu T, Sashinami H, Nakane A, Yamamoto T.** 2003. Lon, a stress-induced ATP-dependent protease, is critically important for systemic *Salmonella enterica* serovar typhimurium infection of mice. Infection and Immunity **71:**690-696.
53. **Bradshaw-Rouse JJ, Whatley MH, Coplin DL, Woods A, Sequeira L, Kelman A.** 1981. Agglutination of *Erwinia stewartii* strains with a corn agglutinin: correlation with extracellular polysaccharide production and pathogenicity. Applied and Environmental Microbiology **42:**344-350.
54. **Duong DA, Stevens AM.** 2017. Integrated downstream regulation by the quorum-sensing controlled transcription factors LrhA and RcsA impacts phenotypic outputs associated with virulence in the phytopathogen *Pantoea stewartii*. PeerJ **5**:e4145.
55. **Dolph PJ, Majerczak DR, Coplin DL.** 1988. Characterization of a gene cluster for exopolysaccharide biosynthesis and virulence in *Erwinia stewartii*. Journal of Bacteriology **170:**865-871.
56. **Grant SG, Jessee J, Bloom FR, Hanahan D.** 1990. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. Proceedings of the National Academy of Sciences of the United States of America **87:**4645-4649.
57. **Kvitko BH, Bruckbauer S, Prucha J, McMillan I, Breland EJ, Lehman S, Mladinich K, Choi KH, Karkhoff-Schweizer R, Schweizer HP.** 2012. A simple method for construction of *pir*+ Enterobacterial hosts for maintenance of R6K replicon plasmids. BMC Research Notes **5:**157.
58. **Simon R, Priefer U, Pühler A.** 1983. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in Gram negative bacteria. Nature Biotechnology **1:**784 - 791.
59. **Labes M, Puhler A, Simon R.** 1990. A new family of RSF1010-derived expression and *lac*-fusion broad-host-range vectors for gram-negative bacteria. Gene **89:**37-46.
60. **Carlier A, Burbank L, von Bodman SB.** 2009. Identification and characterization of three novel EsaI/EsaR quorum-sensing controlled stewartan exopolysaccharide biosynthetic genes in *Pantoea stewartii* ssp. *stewartii*. Molecular Microbiology **74:**903-913.
61. **Stabb EV, Ruby EG.** 2002. RP4-based plasmids for conjugation between *Escherichia coli* and members of the *Vibrionaceae*. Methods in Enzymology **358:**413-426.

Table 5.1. Strains and plasmids used in this study

Strains	Genotype and notes[a]	References
***Pantoea stewartii* strains**		
DC283	Wild-type strain; Nal^r	(55)
Δ*lrhA*	Unmarked deletion of *lrhA* coding sequence from DC283; Nal^r	(18)
Δ*lrhA/lrhA*$^+$	Δ*lrhA* with chromosomal complementation of *lrhA* and its promoter downstream of *glmS*; Nal^r Cm^r	(18)
Δ*rcsA*-2017	Unmarked deletion of *rcsA* coding sequence from DC283; Nal^r	(54)
Δ*rcsA/rcsA*$^+$-2017	Δ*rcsA* with chromosomal complementation of *rcsA* and its promoter downstream of *glmS*; Nal^r Cm^r	(54)
Δ*ompC*	Unmarked deletion of both *ompC* coding sequence from DC283; Nal^r	This study
Δ*ompC/ompC*$^+$	Δ*ompC* with chromosomal complementation of *ompC* and its promoter downstream of *glmS*; Nal^r Cm^r	This study
Δ*ompA*	Unmarked deletion of *ompA* coding sequence from DC283; Nal^r	This study
Δ*ompA/ompA*$^+$	Δ*ompA* with chromosomal complementation of *ompA* and its promoter downstream of *glmS*; Nal^r Cm^r	This study
Δ*lon*	Unmarked deletion of *lon* coding sequence from DC283; Nal^r	This study
Δ*lon/lon*$^+$	Δ*lon* with chromosomal complementation of *lon* and its promoter downstream of *glmS*; Nal^r Cm^r	This study
***Escherichia coli* strains**		
Top 10	*F- mcrA* Δ(*mrr-hsdRMS-mcrBC*) *Φ80dlacZΔM15* Δ*lacX*74 *deoR recAI araD139* Δ(*ara-leu)7697 galU galK rpsL (*Str^r*) endA1 nupG*	(56)
DH5α *λpir*	*F- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169 hsdR17(rK- mK+) λpir*	(57)
S17-1	*recA pro hsdR RP4-2-Tc::Mu-Km::Tn7*	(58)
S17-1 *λpir*	*recA pro hsdR RP4-2-Tc::Mu-Km::Tn7 λpir*	(59)
Plasmids		
pGEM-T	Cloning vector, Ap^r	Promega
pET28a	Expression vector, Kn^r	Novagen
pDONR201	Entry vector in the Gateway system, Kn^r	Life Technologies
pAUC40	Suicide vector pKNG101::*attR-ccdB*-Cm^R; Cm^r, Str^r, *sacB*	(60)
pEVS104	Conjugative helper plasmid, *tra trb;* Kn^r	(61)
pUC18R6K-mini-Tn7-cat	Tn7 vector for chromosomal integration into the intergenic region downstream of *glmS*; Cm^r, Ap^r	(29)
pSAM-DKm	Tn-Seq vector*;* Kn^r	(26)

[a] Ap^{r}, ampicillin resistance; Nal^{r}, nalidixic acid resistance; Kn^{r}, kanamacyin resistance; Cm^{r}, chloramphenicol resistance; Str^{r}, streptomycin resistance.

Table 5.2. Primers used in this study

Deletion construction		
OMPC-UPF	GTCGACATCTCGCCGAATTCTGGA	Amplify 1 kb region upstream of *ompC*
OMPC-UPR	AGTGGAATATAGGCGGCCGCCAGAATG TTGCGCTTCATC	
OMPC-DNF	GCGGCCGCCTATATTCCACTCCAGTTCTA AGTTTGTTGATGC	Amplify 1 kb region downstream of *ompC*
OMPC-DNR	GGATCCACGCGTTCATTATACAGACC	
OMPC-1kbUPF-attB1	GGGGACAAGTTTGTACAAAAAAGCAGG CTGTCGACATCTCGCCGAATTCTGGA	Amplify 2 kb deletion fragment of *ompC* with flanking *attB* sites
OMPC-1kbDNR-attB2	GGGGACCACTTTGTACAAGAAAGCTGG GTGGATCCACGCGTTCATTATACAGACC	
UP- OMPC-F	GGGATGGAAAGATGCGTGCAGAC	Screen/sequence mutants for *ompC* deletion
IN- OMPC-F	ATCTTACACCTCGCCAATCGGC	
DN- OMPC-R	ATTGTCCGGTCGGCAGAATTACC	
OMPA-UPF	GTCGACCTTTGGCGTCAACATTATGC	Amplify 1 kb region upstream of *ompA*
OMPA-UPR	AGTGGAATATAGGCGGCCGCCACTGCA ATTGCGATAGC	
OMPA-DNF	GCGGCCGCCTATATTCCACTCCTCAGGCT TAAGTTATACGTG	Amplify 1 kb region downstream of *ompA*
OMPA-DNR	GGATCCCGATCAGCGAGAGATAATCG	
OMPA-1kbUPF-attB1	GGGGACAAGTTTGTACAAAAAAGCAGG CTGTCGACCTTTGGCGTCAACATTATGC	Amplify 2 kb deletion fragment of *ompA* with flanking *attB* sites
OMPA-1kbDNR-attB2	GGGGACCACTTTGTACAAGAAAGCTGG GTGGATCCCGATCAGCGAGAGATAATCG	
UP- OMPA-F	GCTCCTGCCAGTAATGACAATGGC	Screen/sequence of mutants for *ompA* deletion
IN- OMPA-F	GCCGTATGCCTAACAAAGGCAACG	
DN- OMPA-R	GCCAGATGTCGACGCTGAAGC	
LON-UPF	GTCGACACTACGATGTGCAGAAAGC	Amplify 1 kb region upstream of *lon*
LON-UPR	AGTGGAATATAGGCGGCCGCACGCTCA GGATTCATAGAG	
LON-DNF	GCGGCCGCCTATATTCCACTGGTATGCAG GTAGCTACC	Amplify 1 kb region downstream of *lon*
LON-DNR	GGATCCGCATCTGCTGCATGTAACC	
LON-1kbUPF-attB1	GGGGACAAGTTTGTACAAAAAAGCAGG CTGTCGACACTACGATGTGCAGAAAGC	Amplify 2 kb deletion fragment of *lon* with flanking *attB* sites
LON-1kbDNR-attB2	GGGGACCACTTTGTACAAGAAAGCTGG GTGGATCCGCATCTGCTGCATGTAACC	
UP- LON-F	GGTTTGACCGCACACCTGC	Screen/sequence of mutants for *lon* deletion
IN- LON-F	GGCATCATCCGTTACTACACGCG	
DN- LON-R	CGTACCCACGCTTCCTGTCG	

Chromosomal complementation construction		
C_OMPC_F_SacI	GAGCTCAAGAGATGTTATCCCGTTCCC	Amplify promoter and coding region of *CKS_0458/CKS_0459*, and screen conjugants
C_OMPC_R_XhoI	CTCGAGCTATTAGAACTGGTAAACCACACCC	
C_OMPA_F_SacI	GAGCTCCCAGTCAGACCAGAAAGTCG	Amplify promoter and coding region of *CKS_5208*, and screen conjugants
C_OMPA_R_XhoI	CTCGAGCTATTAAGCCTGAGGCTGAGTTACC	
C_LON_F_SpeI	ACTAGTATGGCAAGCACGAAGCTCAGG	Amplify promoter and coding region of *CKS_5211*, and screen conjugants
C_LON_R_XhoI	CTCGAGCTACTATTTTGCGGTAGCTACCTGC	

Table 5.3. Top candidate genes important for *in planta* growth of *P. stewartii*[a]

Locus_tag[b]	Annotation name[b]	Ratio1[c]	Ratio2[c]
DSJ_10435	MdoG	0.00	0.00
DSJ_09635	porin OmpC	0.00	0.00
DSJ_17215	colicin V production protein	0.00	0.00
DSJ_09680	glycosyl transferase	0.01	0.00
DSJ_09745	porin OmpA	0.01	0.00
DSJ_07830	exopolyphosphatase	0.01	0.02
DSJ_02405	ribonuclease PH	0.01	0.01
DSJ_03910	ribosomal subunit interface protein	0.01	0.01
DSJ_10540	anti-sigma-28 factor FlgM	0.01	0.01
DSJ_08680	L,D-transpeptidase	0.01	0.02
DSJ_19160	endopeptidase La	0.01	0.01
DSJ_07805	phosphoribosylglycinamide formyltransferase	0.01	0.01
DSJ_16055	amylovoran biosynthesis protein AmsB	0.01	0.01
DSJ_00305	shikimate dehydrogenase	0.01	0.01
DSJ_05305	peptidase	0.01	0.01
DSJ_10440	glucan biosynthesis glucosyltransferase H	0.02	0.00
DSJ_00740	hypothetical protein	0.02	0.00
DSJ_22215	ATP-dependent RNA helicase RhlB	0.02	0.02
DSJ_21615	tRNA (adenosine(37)-N6)-dimethylallyltransferaseMiaA	0.02	0.02
DSJ_18350	PTS glucose transporter subunit IIA	0.02	0.01
DSJ_10795	sensor protein PhoQ	0.03	0.01
DSJ_03430	SgbH	0.03	0.01
DSJ_03865	SspA	0.03	0.02
DSJ_00315	hypothetical protein	0.03	0.01
DSJ_04830	hypothetical protein	0.03	0.06
DSJ_04470	DNA repair protein RadA	0.03	0.01
DSJ_17620	hypothetical protein	0.03	0.00
DSJ_08530	kinase inhibitor	0.03	0.02
DSJ_04095	lipoprotein NlpI	0.03	0.02
DSJ_01080	rhomboid family intramembrane serine protease GlpG	0.04	0.01
DSJ_03780	aspartate ammonia-lyase	0.04	0.01
DSJ_19710	hypothetical protein	0.05	0.05
DSJ_00995	GMP/IMP nucleotidase	0.05	0.03
DSJ_15515	phosphate starvation protein PhoH	0.05	0.07
DSJ_10450	SecY/SecA suppressor protein	0.06	0.07
DSJ_22095	adenylate cyclase	0.06	0.02
DSJ_08295	citrate (Si)-synthase	0.06	0.02
DSJ_03785	FxsA	0.06	0.02

DSJ_04620	hypothetical protein	0.07	0.02
DSJ_05575	hypothetical protein	0.07	0.04
DSJ_21690	hypothetical protein	0.07	0.05
DSJ_00170	bifunctional GTP diphosphokinase/guanosine-3',5'-bis(diphosphate) 3'-diphosphatase	0.07	0.03
DSJ_15525	iron uptake system protein EfeO	0.07	0.05
DSJ_04235	ribonuclease E inhibitor B	0.07	0.04
DSJ_03050	phosphate-starvation-inducible protein PsiE	0.08	0.02
DSJ_13265	hypothetical protein	0.08	0.03
DSJ_01165	hypothetical protein	0.10	0.07

[a] Genes with >100 read counts (RC) in LB and <10 RC in *in planta* growth; [b] Locus tag and annotation names are from the complete *P. stewartii* DC283 genome (NCBI accession: CP017581); [c] ratio of number of RC *in planta* growth over number of RC in LB medium

Table 5.4. Genes with between 10 and 100 insertions in LB growth and no insertion mutant retrieved from *in planta* growth[a]

Locus_tag[b]	Annotation name[b]	RC_LB-1[c]	RC_LB-2[c]
DSJ_02860	IS630 family transposase	18	15
DSJ_03320	iron transporter	34.5	47
DSJ_03905	PTS IIA-like nitrogen regulatory protein PtsN	11	28
DSJ_04770	Xaa-Pro aminopeptidase	39	48
DSJ_06075	KsgA	74	131.5
DSJ_06085	peptidylprolyl isomerase SurA	37	50
DSJ_06365	transcriptional regulator PdhR	22	22
DSJ_06395	bifunctional aconitate hydratase 2/2-methylisocitrate dehydratase	14	22
DSJ_06860	thioredoxin TrxC	13	13
DSJ_08165	N-acetylglucosamine-6-phosphate deacetylase	16	15
DSJ_08310	succinate dehydrogenase flavoprotein subunit	10	15
DSJ_08375	cell envelope integrity protein TolA	17	27
DSJ_09685	flippase	65	127
DSJ_10155	hypothetical protein	11	11.5
DSJ_10665	FabG	13	11
DSJ_16045	amylovoran biosynthesis protein AmsD	75	128
DSJ_16070	protein tyrosine phosphatase	11.5	17
DSJ_16855	hypothetical protein	12	21
DSJ_17020	NADH-quinone oxidoreductase subunit M	62	75
DSJ_17060	NADH-quinone oxidoreductase subunit E	17	26
DSJ_17075	NADH-quinone oxidoreductase subunit A	11	23
DSJ_18780	hypothetical protein	12	6
DSJ_18870	phage tail protein	86	115
DSJ_21300	osmotically-inducible protein OsmY	13	16
DSJ_21680	malate dehydrogenase	26	38
DSJ_22350	MarR family transcriptional regulator	50	86

[a] Genes with 10 < read counts (RC) < 100 in LB and zero RC in *in planta* growth; [b] Locus tag and annotation names are from the complete *P. stewartii* DC283 genome (NCBI accession: CP017581); [c] RC in LB samples

Table 5.5. Putative genes important for *in planta* growth of *P. stewartii* in pDSJ010[a]

Locus_tag[b]	Annotation name[b]	Ratio1[c]	Ratio2[c]
DSJ_26225	hypothetical protein	0.00	0.10
DSJ_25990	hypothetical protein	0.00	0.00
DSJ_26100	hypothetical protein	0.00	0.00
DSJ_25505	hypothetical protein	0.05	0.00
DSJ_26120	hypothetical protein	0.05	0.08
DSJ_26000	hypothetical protein	0.07	0.01
DSJ_25285	GlnQ	0.08	0.06
DSJ_25885	branched-chain amino acid ABC transporter permease	0.09	0.04
DSJ_26040	CDP-diacylglycerol diphosphatase	0.09	0.06
DSJ_25895	thiamine biosynthesis protein thio	0.09	0.09

[a] Genes with >10 read counts (RC) in LB and ≥ ten-fold reduction of RC in *in planta* growth;
[b] Locus tag and annotation names are from pDSJ010 of the complete *P. stewartii* DC283 genome (NCBI accession: CP017592); [c] ratio of number of RC *in planta* growth over number of RC in LB medium

Table 5.6. Putative genes with higher fitness *in planta* found by Tn-Seq analysis[a]

Locus_tag[b]	Annotation name[b]	Ratio1[c]	Ratio2[c]
DSJ_16945	DNA-binding response regulator RcsB	53.26	43.56
DSJ_16940	phosphotransferase RcsD	29.03	20.98
DSJ_11065	hypothetical protein	25.68	45.90
DSJ_12140	hypothetical protein	19.47	13.62
DSJ_11900	helix-turn-helix transcriptional regulator RcsA	18.97	46.79
DSJ_16950	two-component system sensor histidine kinase/response regulator RcsC	13.81	17.63
DSJ_12645	hypothetical protein	11.15	13.48

[a] Genes with >10 read counts (RC) in LB and ≥ ten-fold increase of RC in *in planta* growth;
[b] Locus tag and annotation names are from the complete *P. stewartii* DC283 genome (NCBI accession: CP017581); [c] ratio of number of RC *in planta* growth over number of RC in LB medium

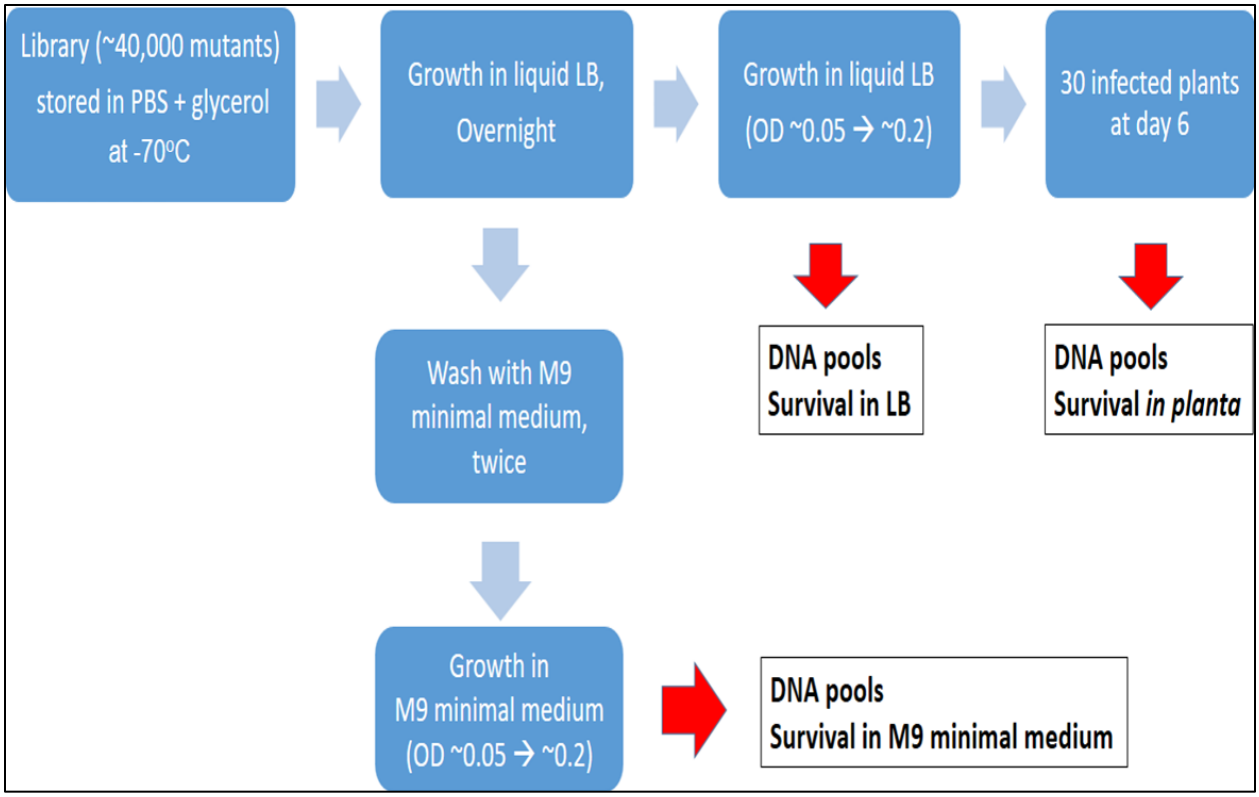

Figure 5.1. Study design to identify virulence factors of *P. stewartii* using Tn-Seq approach.

See text for details. Experiments were done in duplicate.

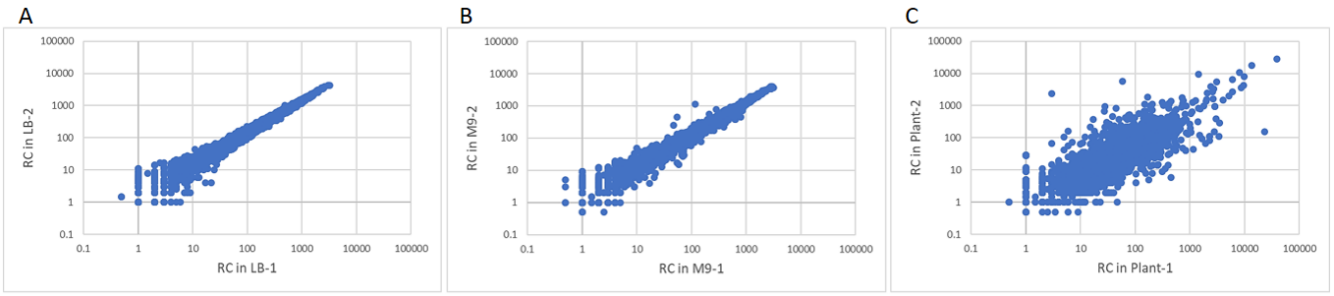

Figure 5.2. The correlation of Tn-Seq data within the same type of samples between the two replicates. A: library grown in LB medium; B: library grown in M9 minimal medium; C: library grown *in planta* samples

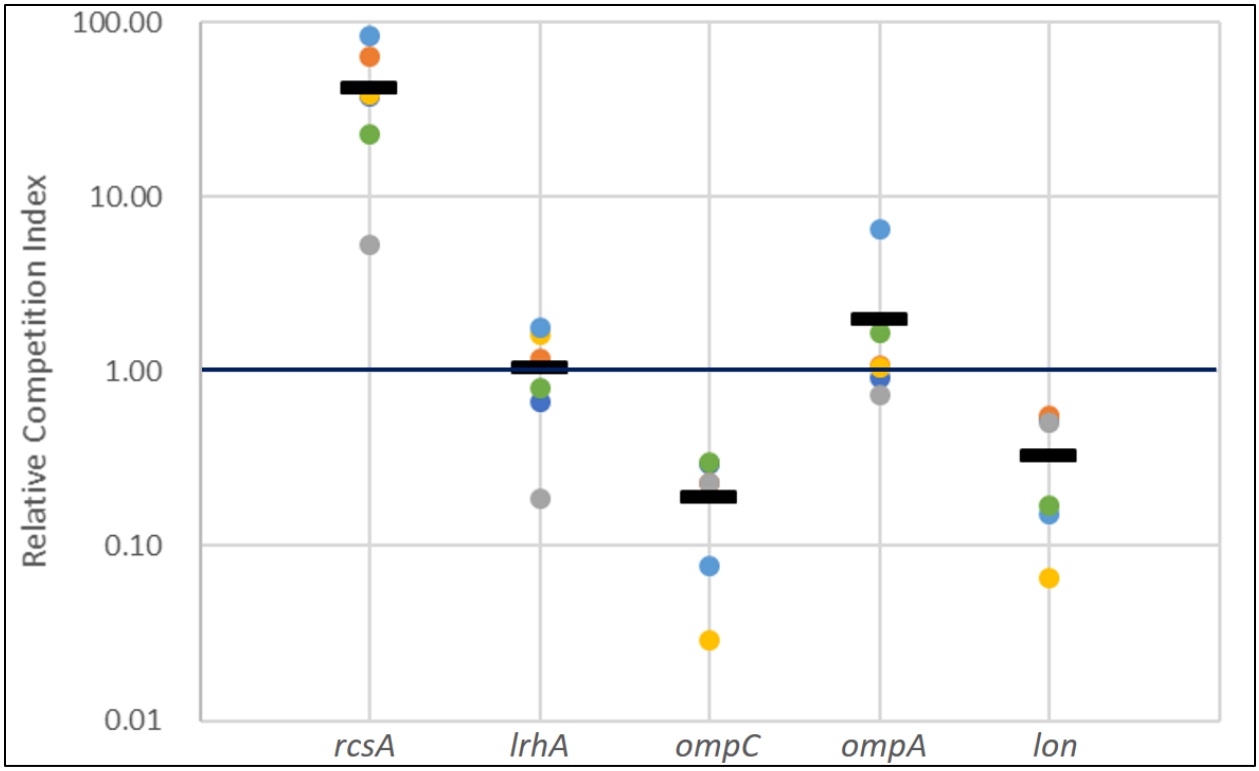

Figure 5.3. Competition assays to validate Tn-Seq analysis. The deletion strains (Nal^R) and complementation strains (Nal^R/Cm^R) were combined in a ~1:1 ratio prior to inoculation into corn seedlings. The competition indices were calculated as the ratios of the deletion strains to the complementation strains after 6 days post-inoculation over the ratios of the deletion strains to the complementation strains of the inoculum.

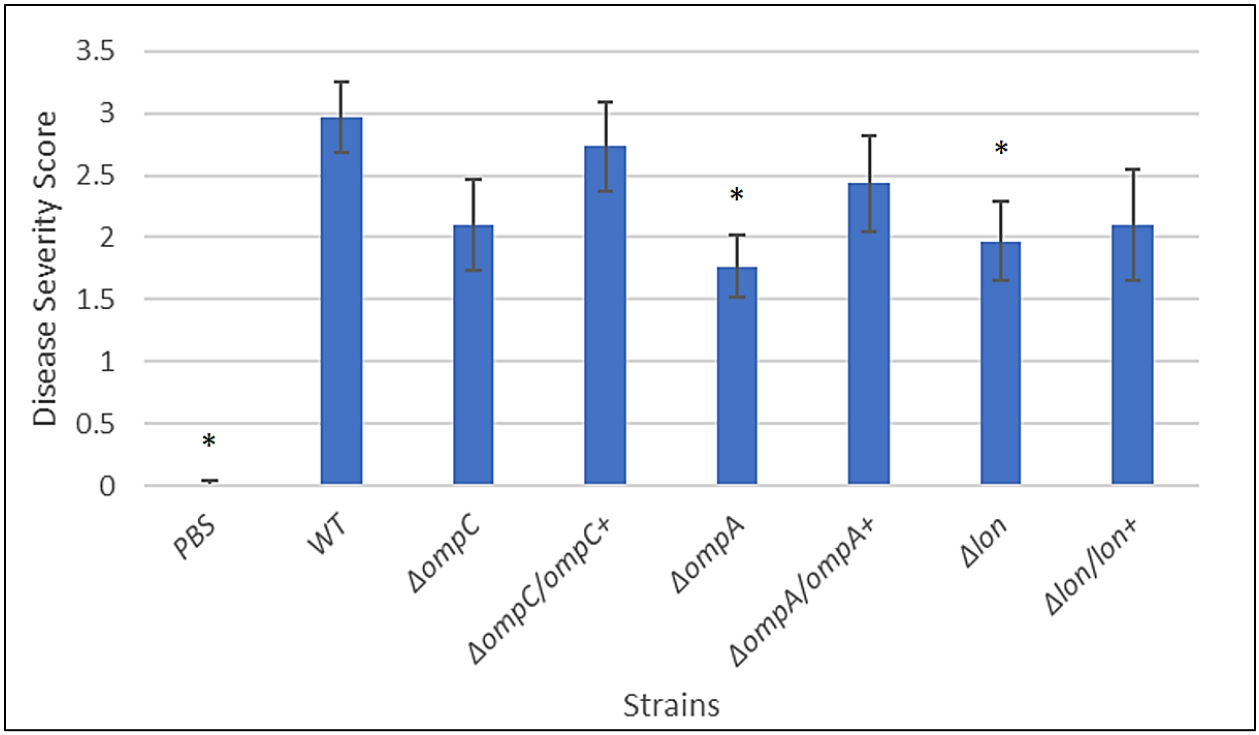

Figure 5.4. Virulence assays of gene candidates for being important for *in planta* survival. The higher disease severity score indicates more severe in the disease symptoms. Error bars denotes standard errors. Asterisk (*) indicates statistically significant difference ($p < 0.05$) between the treatment and wild-type strain using Student's T-test. The observed difference between wild type and *ompC* deletion (Δ*ompC*) strain has a $p = 0.057$.

Chapter Six

Overall Conclusions

Pantoea stewartii subspecies *stewartii* can colonize and infect corn, after being transmitted by the corn flea beetle, *Chaetocnema pulicaria* (1). This results in the development of Stewart's wilt, which is a serious bacterial infection in this crop reducing yield (2, 3). The bacterium *P. stewartii* is an attractive model for studying gene regulation by quorum sensing (QS), bacterial surface motility, and host-pathogen interactions for several reasons. First, the QS master regulator of *P. stewartii*, EsaR, has a unique mechanism of function (4, 5) in comparison with other typical LuxR homologues (6). Second, this bacterium has two hosts, corn and the corn flea beetle, making its life-style interesting to study as well as providing the opportunity to understand the distinct bacterial regulation occurring during cross-kingdom colonization. Third, *P. stewartii* is a risk-group category one organism which is ideal for genetic manipulation since many genetic tools developed for *Escherichia coli* can be used in this organism. Fourth, this phytopathogen also displays surface motility (7), which may contribute to its migration in the plant xylem vessels leading to the formation of biofilm. Surface motility occurs due to multiple factors, including flagellar-mediated movement, type-IV pili, and/or bio-surfactant production. Finally, it has been thought that *P. stewartii* possesses a minimal arsenal of virulence factors to cause disease (2, 8); therefore, it should be simpler to decipher the interaction between this phytopathogen with its host during disease development.

Previous studies have identified the critical role of QS-controlled exopolysaccharide (EPS) production and surface motility in the virulence of *P. stewartii* within corn (2, 7, 9-11). Using global approaches like proteomics and transcriptomics, Ramachandran et. al. systematically identified the QS regulon and defined some direct targets of EsaR (12, 13). Two EsaR direct targets, RcsA and LrhA are transcription factors controlling capsule production and surface motility, respectively, and both of them contribute to the virulence of *P. stewartii* (14).

Research shown in Chapter Two of this dissertation has utilized transcriptomics to in turn define the global regulon downstream of these QS-controlled transcription factors. This has helped to reconfirm the role of RcsA in capsule production as well as to extend the knowledge about the function of LrhA in this phytopathogen. In addition, a regulatory linkage between LrhA and RcsA was identified by the RNA-Seq investigation and confirmed with qRT-PCR validation as well as GFP-transcriptional fusion assays. These findings have provided insights into the hierarchy of QS-controlled network causing wilt disease in corn by *P. stewartii*.

P. stewartii DC283, a nalidixic acid resistant mutant of the original 1976 isolate SS104 from *Zea mays*, is the wild-type reference strain used to study pathogenesis (15, 16). An incomplete draft assembly of the DC283 genome (NCBI GenBank: NZ_AHIE00000000.1) with 65 contigs was released in 2012. Nine more draft *P. stewartii* genomes from many different sources have been released but none have been fully assembled. Chapter Three of this dissertation describes a successful effort to complete the assembly of the *P. stewartii* DC283 genome (17). It resulted in the full genome assembly of this pathogen with one circular chromosome, ten circular plasmids and one linear phage-plasmid. More importantly, the two Type III secretion systems in *P. stewartii* DC283 are now known to be located on two separate mega-plasmids due to the separation of the plasmid sequences from chromosomal DNA. Intriguingly, a novel 66-kb DNA region has been identified in this chromosome which led to the reexamination of the contribution of RcsA to surface motility, as discussed in Chapter Four. In conclusion, the full assembly of a complex prokaryotic genome that contains large numbers of repetitive sequences and many plasmids was achieved as part of this dissertation. The complete genome sequence was immediately utilized to complete an investigation of genes important for *in planta* growth of *P. stewartii* via Tn-Seq (Chapter Five).

In Chapter Two, both transcription factors RcsA and LrhA were shown to play a role in plant virulence and a coordination between LrhA and RcsA regulation was discovered (14). Chapter Four of this dissertation continued to further define the downstream direct targets of LrhA and its interaction with RcsA (18). The existence of a negative autoregulation of LrhA in *P. stewartii* was revealed and confirmed by electrophoretic mobility shift assays (EMSA) using purified LrhA. This is opposite to the positive autoregulation of LrhA found in *E. coli*. In addition, EMSA experiments showed that LrhA binds to the promoter for the RcsA gene, and those for putative fimbrial subunits and biosurfactant production enzymes in *P. stewartii*. However, LrhA could not bind the putative *flhDC* promoter in *P. stewartii.* Since that promoter is a main direct target of LrhA in *E. coli*, this suggests a very distinct regulon of LrhA in *P. stewartii* from its counterpart in *E. coli.* A reexamination of RNA-Seq data generated from Chapter Two using the complete genome from Chapter Three as the reference resulted in the identification of the missing of the 66-kb DNA region in the *rcsA* deletion mutant in Chapter Two. This led to the regeneration of a new *rcsA* deletion mutant and its complementation strain to reexamine the physiological function of RcsA in *P. stewartii*. This study resulted in the confirmation of the role of RcsA in capsule production and plant virulence, but also permitted the new discovery of its role in surface motility. These findings continue to extend our understanding of the coordinated regulatory cascades of QS-controlled regulators utilized in the phytopathogen *P. stewartii*.

Quorum sensing plays a critical role in the virulence of *P. stewartii* through the control of capsule production, surface motility, and other processes. However, other bacterial genes may also contribute to the interactions between *P. stewartii* and corn by promoting successful *in planta* colonization and survival. Therefore, Chapter Five utilized a Tn-Seq global investigation

approach to identify additional bacterial genes essential for growth and survival *in planta*. The Tn-Seq analysis, facilitated by the complete assembly of the *P. stewartii* genome (17), has revealed these important genes by comparing a transposon mutant library grown *in vitro* vs *in planta*. The library, containing ~40,000 mutants, was generated via conjugation that transferred a suicide vector carrying a mariner transposon, pSAM-DKm (19), into the wild-type *P. stewartii* DC283. The genomic DNA extracted from the library grown in LB medium, in M9 minimal medium and *in planta* were collected to conduct massive parallel high-throughput sequencing. The Tn-Seq analysis of the library grown on LB suggested a high number of genes (> 500 genes) are essential or partially essential for growth in the rich medium such as LB. The comparison between the library grown in LB vs in M9 minimal medium, after excluding essential genes for growth in LB, showed that this pathogen required an additional 78 genes to grow well in this minimal medium, using a two-fold cutoff. Many of these genes were involved in the metabolism of purine and amino acids.

The main focus of Chapter Five was the comparison between the transposon mutant library grown in LB vs *in planta* growth. This comparison resulted in several gene categories: (a) 193 genes with more than a ten-fold reduction *in planta* samples with each gene having ≥ 100 read counts in the LB Tn-Seq data; (b) 184 genes with more than a ten-fold reduction *in planta* samples with each gene having between 10-100 read counts in LB Tn-Seq data; (c) 7 genes with greater than ten-fold increase of read counts *in planta* compared to LB samples. Three genes, *ompA, ompC* and *lon*, from the top of the first list of 193 genes were selected to make genetic deletion strains and the corresponding complementation strains in order to validate the Tn-Seq analysis and re-examine their physiological role *in planta*. *In planta* competition assays confirmed the Tn-Seq predictions for *ompC* and *lon* as well as some level of virulence reduction

of these validated genes, presumably due, at least partially, to growth defects. Surprisingly, the competition assays did not suggest a role for *ompA* during *in planta* growth, but OmpA was found to play a role in *P. stewartii* virulence. The mechanism by which *P. stewartii* OmpA contributes to the pathogenicity needs to be further explored. Moreover, the role of additional genes identified by Tn-Seq analysis during *in planta* competition with regard to pathogenesis and survival may be defined through future studies. The findings described in Chapter Five demonstrate the capacity of a Tn-Seq analysis to broaden our understanding of the host-microbe interactions on a genomic-wide scale.

Overall, these projects have advanced the knowledge about QS controlled transcription regulators in the phytopathogen *P. stewartii*. The complete assembly of the *P. stewartii* genome is an important landmark to further investigation of host-microbe interactions using high-throughput sequencing techniques. A pioneering effort of utilization of Tn-Seq analysis has been validated to be applied in this organism. All of these findings open new doors for researchers in the field to explore and comprehend the biological interactions occurring in the system as well as to invent new approaches for control, prevention and elimination of the disease. For example, some genes in the novel 66-kb region of *P. stewartii* genome are contributing factors to the virulence and surface motility of the pathogen. Identification of these genes will provide additional insights into the mechanism of surface motility utilized by this bacterium. Tn-Seq analysis on this region suggested that ten genes may contribute to the *in planta* growth of the pathogen, in which DSJ_09180 (encoding a transcriptional regulator MntR), DSJ_09315 (encoding a sensory transduction regulator) and DSJ_09140 (encoding a 23S rRNA (adenine(1618)-N(6))-methyltransferase) may be worthy of further investigation.

Approaches similar to those used in this dissertation, including RNA-Seq and Tn-Seq, can be applied to expand our understanding about this phytopathogen in the interaction with corn at an early stage of the infection using a leaf infection assay (in collaboration with David Mackey at the Ohio State University) and with corn flea beetles (in collaboration with Thomas Kuhar at Virginia Tech). On the other hand, research in using quorum quenching (QQ) to interrupt the bacterial infection in corn should be pursued since QS plays such a prominent role in the pathogenicity of *P. stewartii*. QQ is a mechanism of QS interference/inhibition which can be achieved by: (a) interruption of the QS signal biosynthesis, (b) blockage of the QS master regulator activity, and (c) degradation of the QS signal prior to its accumulation and activation of the regulator (20-22). In conclusion, the expanded knowledge about *P. stewartii* pathogenesis and microbe-host interactions gained from this dissertation helps lay the groundwork for the development of future disease prevention strategies.

REFERENCES

1. **Esker PD, Nutter FW.** 2002. Assessing the risk of Stewart's disease of corn through improved knowledge of the role of the corn flea beetle vector. Phytopathology **92:**668-670.
2. **Roper MC.** 2011. *Pantoea stewartii* subsp. *stewartii*: lessons learned from a xylem-dwelling pathogen of sweet corn. Molecular Plant Pathology **12:**628-637.
3. **Tambong JT.** 2015. Specific identification and detection of *Pantoea stewartii* subsp. *stewartii* using a membrane-based multi-gene oligonucleotide array. Canadian Journal of Plant Pathology **37** 414-426.
4. **Minogue TD, Wehland-von Trebra M, Bernhard F, von Bodman SB.** 2002. The autoregulatory role of EsaR, a quorum-sensing regulator in *Pantoea stewartii* ssp. *stewartii*: evidence for a repressor function. Molecular Microbiology **44:**1625-1635.
5. **Schu DJ, Carlier AL, Jamison KP, von Bodman S, Stevens AM.** 2009. Structure/function analysis of the *Pantoea stewartii* quorum-sensing regulator EsaR as an activator of transcription. Journal of Bacteriology **191:**7402-7409.
6. **Stevens AM, Queneau Y, Soulere L, von Bodman S, Doutheau A.** 2011. Mechanisms and synthetic modulators of AHL-dependent gene regulation. Chemical Reviews **111:**4-27.
7. **Herrera CM, Koutsoudis MD, Wang X, von Bodman SB.** 2008. *Pantoea stewartii* subsp. *stewartii* exhibits surface motility, which is a critical aspect of Stewart's wilt disease development on maize. Molecular Plant-Microbe Interactions **21:**1359-1370.
8. **Braun EJ.** 1982. Ultrastructural investigation of resistant and susceptible maize inbreds infected with *Erwinia stewartii*. Phytopathology **72:**159-166.
9. **Bradshaw-Rouse JJ, Whatley MH, Coplin DL, Woods A, Sequeira L, Kelman A.** 1981. Agglutination of *Erwinia stewartii s*trains with a corn agglutinin: correlation with extracellular polysaccharide production and pathogenicity. Applied and Environmental Microbiology **42:**344-350.
10. **von Bodman SB, Majerczak DR, Coplin DL.** 1998. A negative regulator mediates quorum-sensing control of exopolysaccharide production in *Pantoea stewartii* subsp. *stewartii*. Proceedings of the National Academy of Sciences of the United States of America **95:**7687-7692.
11. **Minogue TD, Carlier AL, Koutsoudis MD, von Bodman SB.** 2005. The cell density-dependent expression of stewartan exopolysaccharide in *Pantoea stewartii* ssp. *stewartii* is a function of EsaR-mediated repression of the *rcsA* gene. Molecular Microbiology **56:**189-203.
12. **Ramachandran R, Stevens AM.** 2013. Proteomic analysis of the quorum-sensing regulon in *Pantoea stewartii* and identification of direct targets of EsaR. Applied and Environmental Microbiology **79:**6244-6252.
13. **Ramachandran R, Burke AK, Cormier G, Jensen RV, Stevens AM.** 2014. Transcriptome-based analysis of the *Pantoea stewartii* quorum-sensing regulon and identification of EsaR direct targets. Applied and Environmental Microbiology **80:**5790-5800.
14. **Kernell Burke A, Duong DA, Jensen RV, Stevens AM.** 2015. Analyzing the transcriptomes of two quorum-sensing controlled transcription factors, RcsA and LrhA, important for *Pantoea stewartii* virulence. PLOS ONE **10:**e0145358.
15. **Coplin DL, Rowan RG, Chisholm DA, Whitmoyer RE.** 1981. Characterization of plasmids in *Erwinia stewartii*. Applied and Environmental Microbiology **42:**599-604.
16. **Coplin DL, Majerczak DR, Zhang Y, Kim W, Jock S, Geider K.** 2002. Identification of *Pantoea stewartii* subsp. *stewartii* by PCR and strain differentiation by PFGE. Plant Disease **86:**304-311.
17. **Duong DA, Stevens AM, Jensen RV.** 2017. Complete genome assembly of *Pantoea stewartii* subsp. *stewartii* DC283, a corn pathogen. Genome Announcements **5**: e00435-17

18. **Duong DA, Stevens AM.** 2017. Integrated downstream regulation by the quorum-sensing controlled transcription factors LrhA and RcsA impacts phenotypic outputs associated with virulence in the phytopathogen *Pantoea stewartii*. PeerJ **5**:e4145.
19. **Gutierrez MG, Yoder-Himes DR, Warawa JM.** 2015. Comprehensive identification of virulence factors required for respiratory melioidosis using Tn-seq mutagenesis. Frontiers in Cellular and Infection Microbiology **5:**78.
20. **Chernin L, Toklikishvili N, Ovadis M, Kim S, Ben-Ari J, Khmel I, Vainstein A.** 2011. Quorum-sensing quenching by rhizobacterial volatiles. Environmental Microbiology Reports **3:**698-704.
21. **Helman Y, Chernin L.** 2015. Silencing the mob: disrupting quorum sensing as a means to fight plant disease. Molecular Plant Pathology **16:**316-329.
22. **Dong YH, Wang LY, Zhang LH.** 2007. Quorum-quenching microbial infections: mechanisms and implications. Philosophical transactions of the Royal Society of London. Series B, Biological Sciences **362:**1201-1211.

APPENDIX A

CHAPTER TWO SUPPLEMENTARY INFORMATION

Table A.1: Primers used for strain construction

Primers	Sequence (5' to 3')	Purpose
Deletion construction		
LrhA-UPF	GTCGACATTGTCCAGTTTGCCGG	Amplify 1 kb region upstream of *lrhA*
LrhA-UPR	AGTGGAATATAGGCGGCCGCCTTCACTTATTAGAG	
LrhA-DNF	GCGGCCGCCTATATTCCACTATCCCGTCTTC	Amplify 1 kb region downstream of *lrhA*
LrhA-DNR	GGATCCCCAATGCGCACCAG	
LrhA-1kbUPF-attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTGTCGACGTTTGCCGGATTTATCAATTTG	Amplify 2 kb deletion fragment of *lrhA* with flanking *attB* sites
LrhA-1kbDNR-attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTGGATCCGCGCACCAGATAAACCAGGC	
UP-LrhA-SeqF	GTATGACAGACCCATTTACCCCG	Screen/sequence mutants for *lrhA* deletion
IN-LrhA-SeqF	GCGATCCCTCTGGTATTGCTGG	
DN-LrhA-SeqR	GCCCTGTTGGCCAGAGTATG	
RcsA-UPF	GTCGACATCCTTCAACGGTCATTTGTG	Amplify 1 kb region upstream of *rcsA*
RcsA-UPR	AGTGGAATATAGGCGGCCGCCTCACCAATTTGTTATC	
RcsA-DNF	GCGGCCGCCTATATTCCACTCCTCACAGAACTG	Amplify 1 kb region downstream of *rcsA*
RcsA-DNR	GGATCCGATAGCGCCTTCAAGC	
RcsA-1kbUPF-attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTGTCGACCAACGGTCATTTGTGGCTTATC	Amplify 2 kb deletion fragment of *rcsA* with flanking *attB* sites
RcsA-1kbDNR-attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTGGATCCGCCTTCAAGCACCGAACCGAC	
UP-RcsA-SeqF	TACCCGATGTGGATTCGACGCC	Screen/sequence of mutants for *rcsA* deletion
IN-RcsA-SeqF	CCAGCCCTGAGGAAATACGCGG	
DN-RcsA-SeqR	TTGAAATCATCCATCTCACCAATACGCGC	
Chromosomal complementation construction		
EcoRI-PlrhA-F	GAATTCTGCACAATGTACTCTCCTCACG	Amplify promoter and coding region of *lrhA*, and screen conjugants
XhoI-LrhA-R	CTCGAGCTATTACTCTTCATCGTCCAGCAG	
SpeI-RcsA-F	ACTAGTGAAATTCACAACTATCCGGGCATTTTTC	Amplify promoter and coding region of *rcsA*, and screen conjugants
SacI-RcsA-R	GAGCTCCTATCTTACGTTGACGTAAATACCAG	
glmS downstream	TCTCTGATAAGCACCATGCCCTGT	PCR screen and sequence for validation of chromosomal insertion [1]
glmS intergenic region rev	TACGGTGCTACGCATCAGTGTCAT	
GFP plasmid construction		
TFRcsAFwd EcoRI	GAATTGGAAATTCACAACTATCCGGGCATTTTTC	Primers used for GFP transcriptional fusion construction
TFRcsARev KpnI	GGTACCGTTAGCGACCCTCACCAATTTGTTATCC	

Table A.2: Primers used for qRT-PCR

Gene	Primer	Annealing Temp	Sequence 5' to 3'
Genes validated for RcsA regulon			
argC	Cloning-Forward	55°C	ATCGTTGGTGCCAGTGGTTACG
	Cloning-Reverse		AGGGAAACCGAAACGAATATTAAGGCAC
	RT-PCR Forward	62°C	GTGGAGCAGGGCGCAAA
	RT-PCR Reverse		AAATACCGTAAGGCTGCAGACTGA
CKS_2806	Cloning-Forward	55°C	GAACGCGGTCTGGAACGG
	Cloning-Reverse		TCAGGCAATGCGTTGGGTG
	RT-PCR Forward	64°C	GCGAAGGCCAGAATGTTGACA
	RT-PCR Reverse		GCTCACGCGACGTGTTACG
CKS_3504	Cloning-Forward	55°C	ATGCCAGCCGAAACATCTGCAGCA
	Cloning-Reverse		TGCCGGGGAGACTGAATGGG
	RT-PCR Forward	64°C	GCGTTTTTCGTGCCATGGA
	RT-PCR Reverse		CACTTTGCCCTGGGTGATCA
wceG2	Previously published [2]		
wza	Cloning-Forward	55°C	ATGATTACAATGAAAATGAAGATGATACC TGTTTTGG
	Cloning-Reverse		TTAGTTCGACCAGTTGCGGATGC
	RT-PCR Forward	62°C	GCGAACAGCGCGTGTCA
	RT-PCR Reverse		ATGGTTTTGGCTCAGAT
Genes validated for LrhA regulon			
CKS0458	Previously published [2]		
CKS_3793	Cloning-Forward	57°C	ATGCTAGATATCGTCGAACTGTC
	Cloning-Reverse		TCTGTTCATGGTGATAGCGC
	RT-PCR Forward	64°C	CCTTTGTGGGCCTGTTCTTTTT
	RT-PCR Reverse		ACCGCCAGATGCTGCACTT
CKS_5208	Cloning-Forward	57°C	ATGAAAGTGATTATTGGCGCAC
	Cloning-Reverse		TCTGTGGGTTTGATTAGCCAG
	RT-PCR Forward	60°C	TCAATAACGAACCGCAGTCGAT
	RT-PCR Reverse		AATCCCTCGCGCGCTTT
CKS_5211	Cloning-Forward	57°C	ATGAGTCACGGTTATTCTGTAGTAAAC
	Cloning-Reverse		TCAACAAAGCCTACCACAGC
	RT-PCR Forward	60°C	CATTACGTTTGATCTTCCGTTCTCA
	RT-PCR Reverse		GACGATATGGCAGCCCTGTTC
rcsA	Cloning-Forward	55°C	ACTAGTGAAATTCACAACTATCCGGGCAT TTTTC
	Cloning-Reverse		GAGCTCCTATCTTACGTTGACGTAAATAC CAG

	RT-PCR Forward	60°C	AGCGGAAAATTAAAACGCACAAC
	RT-PCR Reverse		CAGAGGTCACGTTATCGGTTAAGC
Gene used for normalization of samples			
16S rRNA	Previously published [2]		

REFERENCES

1. **Burbank L, Roper MC.** 2014. OxyR and SoxR modulate the inducible oxidative stress response and are implicated during different stages of infection for the bacterial phytopathogen *Pantoea stewartii* subsp. *stewartii*. Molecular Plant-Microbe Interactions **27:**479-490.
2. **Ramachandran R, Burke AK, Cormier G, Jensen RV, Stevens AM.** 2014. Transcriptome-based analysis of the *Pantoea stewartii* quorum-sensing regulon and identification of EsaR direct targets. Applied and Environmental Microbiology **80:**5790-5800.

APPENDIX B

CHAPTER FOUR SUPPLEMENTARY INFORMATION

Table B.1. Primers used for strain construction.

Primers	Sequence (5' to 3')	Purpose
GFP transcriptional fusion		
EcoRI-PlrhA-F	GAATTCTGCACAATGTACTCTCCTCAC G	GFP transcriptional fusion construction
KpnI- PlrhA-R	GGTACCGAGTTTTAAGAACTACTATCC TGG	
LrhA overexpression		
BamHI-LrhA-F	GGATCCATGACTAATGCAAATCGTCCG	*lrhA* coding sequence cloning for overexpression
HindIII-LrhA-R	AAGCTTCTATTACTCTTCATCGTCCAG CAG	
Targets of LrhA for EMSA		
FAM_PlrhA_R	/FAM/GAGTTTTAAGAACTACTATCCTG G	Generation of P_{lrhA}-FAM probe
SpeI-LrhA-F	ACTAGTCATAGCGTAAGTAGGGTGTGA C	
FAM_PrcsA_R	/FAM/GTTAGCGACCCTCACCAATTTGT TATCC	Generation of P_{rcsA}-FAM probe
TFRcsAFwd EcoRI (1)	GAATTGGAAATTCACAACTATCCGGGC ATTTTTC	
PflhDC-R-FAM	/FAM/ATCCTGAGAGACGCGACGTAAC GG	Generation of P_{flhDC}-FAM probe
PflhDC-F2	GACCGGACAGACAGGGTCGC	
P5211-R-FAM	/FAM/CGTGACTCATATCACTCTCCCCA CC	Generation of P_{CKS_5211}-FAM probe
P5211-F1	CGGTGTTTCACGCGATAGTGTCC	
P5208-R-FAM	/FAM/CGCCAATAATCACTTTCATCTCA GACACC	Generation of P_{CKS_5208}-FAM probe
P5208-F	GCTGTGGTAGGCTTTGTTGAATAAATG G	
P0458-R-FAM	/FAM/TAGCAACACCGGCAGCTAACAG	Generation of P_{CKS_0458}-FAM probe
P0458-F	GATTTCCGAATAGCGCCGAACACG	
BamHI- PlrhA-R	GGATCCGAGTTTTAAGAACTACTATCC TGG	Generation of P_{lrhA} competitor
SpeI-LrhA-F	ACTAGTCATAGCGTAAGTAGGGTGTGA C	
Deletion construction		
CKS_0458/CKS_0459-UPF	GTCGACATCGCGACAATGCCGCAGC	Amplify 1 kb region upstream of *CKS_0458/CKS_0459*
CKS_0458/CKS_0459-UPR	AGTGGAATATAGGCGGCCGCAAGCTT GAATTTCATAACC	

CKS_0458/CKS_0459-DNF	GCGGCCGCCTATATTCCACTGTCACGCTGACCTGGTGA	Amplify 1 kb region downstream of *CKS_0458/CKS_0459*
CKS_0458/CKS_0459-DNR	GGATCCCGTTGTAAGTCGCCTGCCCGC	
CKS_0458/CKS_0459-1kbUPF-attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTGTCGACATCGCGACAATGCCGCAGC	Amplify 2 kb deletion fragment of *CKS_0458/CKS_0459* with flanking *attB* sites
CKS_0458/CKS_0459-1kbDNR-attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTGGATCCCGTTGTAAGTCGCCTGCC	
UP-CKS_0458/CKS_0459-SeqF	GGCGCTGGAGAGAGTGACGG	Screen/sequence mutants for *CKS_0458/CKS_0459* deletion
IN-CKS_0458/CKS_0459-SeqF	GTGCCTGTGTTGGCGTAGATGC	
DN-CKS_0458/CKS_0459-SeqR	CCCTGAGCCTGAGGCAACACC	
CKS_5208-UPF	GTCGACGAGAGTGATATGAGTCACGG	Amplify 1 kb region upstream of *CKS_5208*
CKS_5208-UPR	AGTGGAATATAGGCGGCCGCCACTTTCATCTCAGACACC	
CKS_5208-DNF	GCGGCCGCCTATATTCCACTTAAACGCGTCTTCAGGC	Amplify 1 kb region downstream of *CKS_5208*
CKS_5208-DNR	GGATCCGATTTCCCGCAGTATTTCTCG	
CKS_5208-1kbUPF-attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTGTCGACGAGAGTGATATGAGTCACGG	Amplify 2 kb deletion fragment of *CKS_5208* with flanking *attB* sites
CKS_5208-1kbDNR-attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTGGATCCGATTTCCCGCAGTATTTCTCG	
UP-CKS_5208-SeqF	ATCTTGAGCAGATTGCCACGC	Screen/sequence of mutants for *CKS_5208* deletion
IN-CKS_5208-SeqF	CTGTGCAACTGGCTAATCAAACCC	
DN-CKS_5208-SeqR	GCGTCACTGGCACAGTATATGG	
CKS_5211-UPF	GTCGACCCAGTTGAACAGGAGATTATCG	Amplify 1 kb region upstream of *CKS_5211*
CKS_5211-UPR	AGTGGAATATAGGCGGCCGCGTTTACTACAGAATAACCGTG	
CKS_5211-DNF	GCGGCCGCCTATATTCCACTGATTTTATGCTGTGGTAGGC	Amplify 1 kb region downstream of *CKS_5211*
CKS_5211-DNR	GGATCCGGGTTCGTATAACACAATCG	
CKS_5211-1kbUPF-attB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTGTCGACCCAGTTGAACAGGAGATTATCG	Amplify 2 kb deletion fragment of *CKS_5211* with flanking *attB* sites
CKS_5211-1kbDNR-attB2	GGGGACCACTTTGTACAAGAAAGCTGGGTGGATCCGGGTTCGTATAACACAAT	

	CG	
UP-CKS_5211-SeqF	ATAAGACCAGCCTCCCTTTCCTCG	Screen/sequence of mutants for *CKS_5211* deletion
IN-CKS_5211-SeqF	GTGTCCCGACCCGTAACAGG	
DN-CKS_5211-SeqR	CCGTCATACGCAAGCATGTAACGC	
Chromosomal complementation construction		
SacI-CKS_0458/CKS_0459-F	GAGCTCAGTGATTTCCGAATAGCGCCG	Amplify promoter and coding region of *CKS_0458/CKS_0459*, and screen conjugants
XhoI-CKS_0458/CKS_0459-R	CTCGAGCTATCACCAGGTCAGCGTGAC	
XhoI- CKS_0458-R	CTCGAGCTACACTTAATAGTTCACGGCAACG	Work with SacI-CKS_0458/CKS_0459-F to amplify promoter and coding region of *CKS_0458*, and screen conjugants
SacI-CKS_5208-F	GAGCTCAACGGCGACCTGGATATGGC	Amplify promoter and coding region of *CKS_5208*, and screen conjugants
XhoI-CKS_5208-R	CTCGAGCTATTAACGGGTCAGAAAGCGTTCC	
SacI-CKS_5211-F	GAGCTCATCGACTGCTCCACCATGG	Amplify promoter and coding region of *CKS_5211*, and screen conjugants
XhoI-CKS_5211-R	CTCGAGCTATCAACAAAGCCTACCACAGC	

Table B.2. List of 68 genes present in the 66-kb deletion region in Δ*rcsA*-2015.

Locus_tag	Annotated Name	Length (bp)	Coding Strand
DSJ_09100	IS630 family transposase	1054	Forward
DSJ_09105	histidine utilization repressor	756	Forward
DSJ_09110	HutD family protein	546	Reverse
DSJ_09115	formimidoylglutamate deiminase	1356	Reverse
DSJ_09120	imidazolonepropionase	1218	Forward
DSJ_09125	N-formylglutamate deformylase	789	Forward
DSJ_09130	Asp/Glu/hydantoin racemase	636	Reverse
DSJ_09135	NAD(P)H-binding protein	1107	Forward
DSJ_09140	23S rRNA (adenine(1618)-N(6))-methyltransferase	924	Forward
DSJ_09145	glnQ	723	Reverse
DSJ_09150	glutamine ABC transporter permease GlnP	660	Reverse
DSJ_09155	glutamine ABC transporter substrate-binding protein GlnH	747	Reverse
DSJ_09160	DNA starvation/stationary phase protection protein Dps	504	Reverse
DSJ_09165	threonine/homoserine exporter RhtA	912	Reverse
DSJ_09170	ompX	513	Forward
DSJ_09175	hypothetical protein	186	Reverse
DSJ_09180	transcriptional regulator MntR	459	Forward
DSJ_09185	anion transporter	1110	Forward
DSJ_09190	multidrug export protein EmrA	1149	Forward
DSJ_09195	EmrB/QacA family drug resistance transporter	1569	Forward
DSJ_09200	RND transporter	1530	Forward
DSJ_09205	DNA-binding protein	402	Forward
DSJ_09210	ABC-F family ATPase	1593	Forward
DSJ_09215	molybdopterin-synthase adenylyltransferase MoeB	768	Reverse
DSJ_09220	molybdopterin molybdotransferase	1236	Reverse
DSJ_09225	beta-aspartyl-peptidase	960	Forward
DSJ_09230	glutathione ABC transporter ATP-binding protein GsiA	1830	Forward
DSJ_09235	glutathione ABC transporter substrate-binding protein GsiB	1536	Forward
DSJ_09240	glutathione ABC transporter permease GsiC	927	Forward
DSJ_09245	glutathione ABC transporter permease GsiD	906	Forward
DSJ_09250	cytoplasmic protein	450	Forward
DSJ_09255	rimO	1329	Reverse
DSJ_09260	carbonic anhydrase	636	Forward
DSJ_09265	oxidoreductase	1137	Forward
DSJ_09270	hydrolase	657	Reverse
DSJ_09275	serine-type D-Ala-D-Ala carboxypeptidase	1203	Forward
DSJ_09280	undecaprenyl-diphosphate phosphatase	606	Reverse
DSJ_09285	transporter	1689	Reverse

DSJ_09290	hypothetical protein	366	Forward
DSJ_09295	glutaredoxin	264	Reverse
DSJ_09300	hypothetical protein	255	Forward
DSJ_09305	nitroreductase A	723	Forward
DSJ_09310	ribosomal protein S6 modification protein	903	Forward
DSJ_09315	sensory transduction regulator	480	Forward
DSJ_09320	IS66 family transposase	1572	Reverse
DSJ_09325	transposase	348	Reverse
DSJ_09330	transposase	630	Reverse
DSJ_09335	spermidine/putrescine ABC transporter substrate-binding protein PotF	1110	Forward
DSJ_09340	potG	1134	Forward
DSJ_09345	putrescine ABC transporter permease PotH	963	Forward
DSJ_09350	putrescine ABC transporter permease PotI	846	Forward
DSJ_09355	hypothetical protein	468	Forward
DSJ_09360	rumB	1146	Forward
DSJ_09365	arginine ABC transporter substrate-binding protein	732	Reverse
DSJ_09370	artM	669	Reverse
DSJ_09375	arginine transporter permease subunit ArtQ	717	Reverse
DSJ_09380	arginine ABC transporter substrate-binding protein	732	Reverse
DSJ_09385	arginine ABC transporter ATP-binding protein ArtP	750	Reverse
DSJ_09390	lipoprotein	540	Reverse
DSJ_09395	hypothetical protein	321	Forward
DSJ_09400	N-acetylmuramoyl-L-alanine amidase	825	Forward
DSJ_09405	low-specificity L-threonine aldolase	1062	Reverse
DSJ_09410	pyruvate oxidase	1722	Reverse
DSJ_09415	hypothetical protein	900	Reverse
DSJ_09420	hypothetical protein	306	Forward
DSJ_09425	ATP-dependent endonuclease	1344	Forward

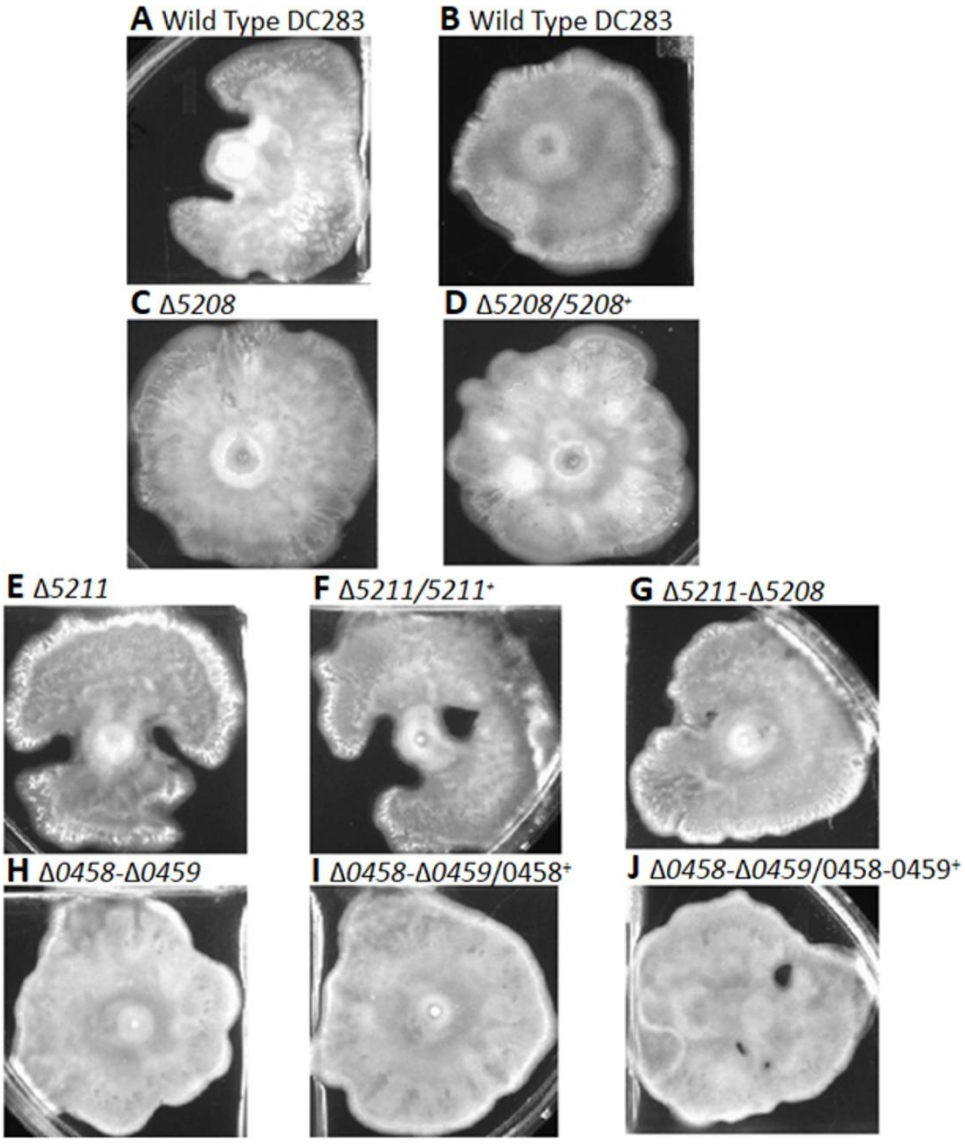

Figure B.1: Impact of putative fimbrial and surfactant genes on surface motility. Surface motility assays for the indicated strains. All pictures were taken at the same magnification after 48 hours of incubation.

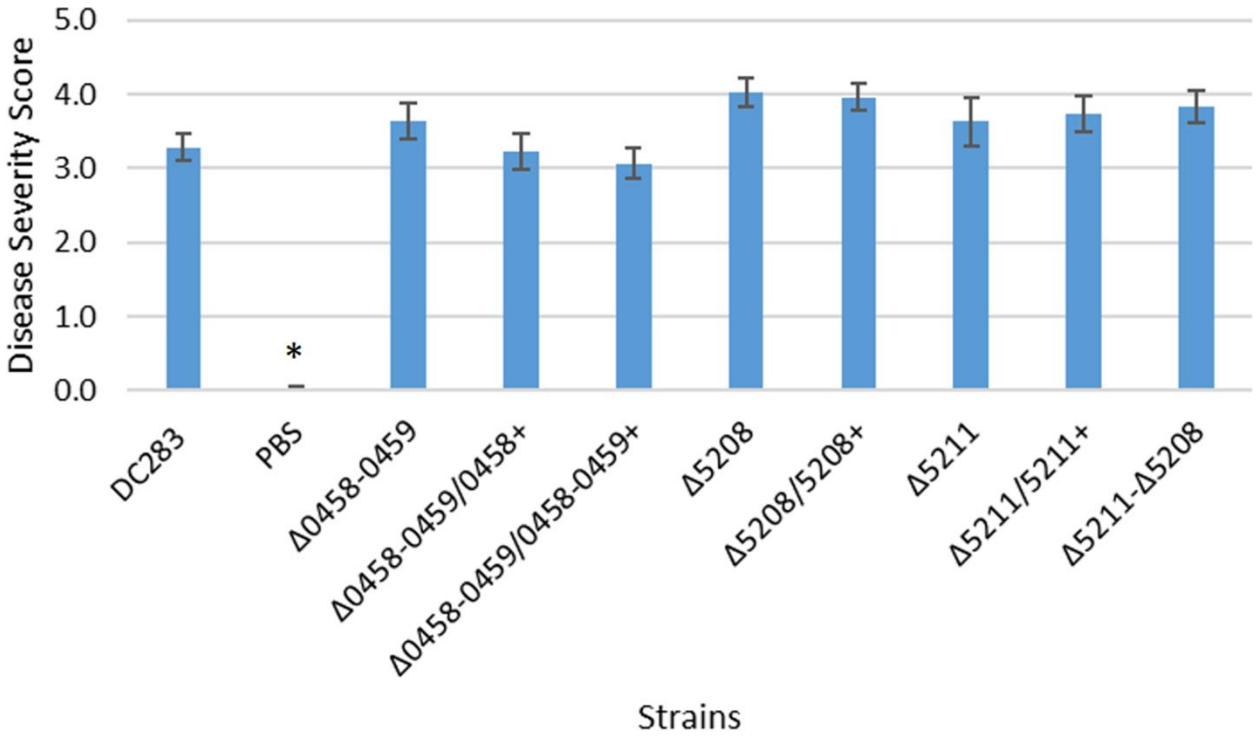

Figure B.2: Xylem-infection assays testing the role of putative fimbrial and surfactant genes in virulence. Data shown is the average score of disease for Day 12 of an infection assay performed with 15 plants inoculated with *P. stewartii* DC283 strains or PBS as a negative control as indicated on the X-axis. Error bars denote standard errors. The asterisk (*) indicates a statistically significant difference ($p<0.05$) between the wild type and the negative control while the remaining strains have $p>0.05$ using a two-tailed homoscedastic Student's t-test.

REFERENCES

1. **Kernell Burke A, Duong DA, Jensen RV, Stevens AM.** 2015. Analyzing the transcriptomes of two quorum-sensing controlled transcription factors, RcsA and LrhA, important for *Pantoea stewartii* virulence. PLOS ONE **10:**e0145358.

APPENDIX C

CHAPTER FIVE SUPPLEMENTARY INFORMATION

Table C.1. Genes with less than ten read counts in library grown in LB medium[a]

Locus_tag[b]	Annotation name[b]	CKS_#[c]	Gene[c]	LB-1[d]	LB-2[d]
DSJ_00005	dnaA	CKS_4348	*dnaA*	0	4
DSJ_00010	DNA polymerase III subunit beta	CKS_4347	*dnaN*	1	1
DSJ_00020	DNA gyrase subunit B	CKS_4345	*gyrB*	2	9
DSJ_00095	DNA replication protein DnaC	CKS_4327	*dnaC*	1	1
DSJ_00100	primosomal protein DnaI	CKS_4326	*dnaT*	1	1
DSJ_00160	guanylate kinase	CKS_4311	*gmk*	1	3
DSJ_00165	**DNA-directed RNA polymerase subunit omega**	**CKS_4310**	***rpoZ***	**0**	**0**
DSJ_00245	YihA family ribosome biogenesis GTP-binding protein	CKS_4292	*yihA*	1	2
DSJ_00350	peptide deformylase	CKS_1554	*def*	2	1
DSJ_00355	methionyl-tRNA formyltransferase	CKS_1555	*fmt*	1	4
DSJ_00385	**50S ribosomal protein L17**	**CKS_1562**	***rplQ***	**0**	**0**
DSJ_00390	DNA-directed RNA polymerase subunit alpha	CKS_1563	*rpoA*	1	3
DSJ_00395	30S ribosomal protein S4	CKS_1564	*rpsD*	0	1
DSJ_00400	**30S ribosomal protein S11**	**CKS_1565**	***rpsK***	**0**	**0**
DSJ_00405	30S ribosomal protein S13	CKS_1566	*rpsM*	1	3
DSJ_00410	**50S ribosomal protein L36**	**CKS_1567**	***rpmJ***	**0**	**0**
DSJ_00415	preprotein translocase subunit SecY	CKS_1568	*secY*	1	3
DSJ_00420	50S ribosomal protein L15	CKS_1569	*rplO*	0	1
DSJ_00425	50S ribosomal protein L30	CKS_1570	*rpmD*	2	0
DSJ_00430	30S ribosomal protein S5	CKS_1571	*rpsE*	0	1
DSJ_00435	**50S ribosomal protein L18**	**CKS_1572**	***rplR***	**0**	**0**
DSJ_00440	50S ribosomal protein L6	CKS_1573	*rplF*	1	3
DSJ_00445	**30S ribosomal protein S8**	**CKS_1574**	***rpsH***	**0**	**0**
DSJ_00450	**rpsN**	**CKS_1575**	***rpsN***	**0**	**0**
DSJ_00455	50S ribosomal protein L5	CKS_1576	*rplE*	1	2
DSJ_00460	50S ribosomal protein L24	CKS_1577	*rplX*	1	1
DSJ_00465	50S ribosomal protein L14	CKS_1578	*rplN*	2	1
DSJ_00470	30S ribosomal protein S17	CKS_1579	*rpsQ*	0	2
DSJ_00475	50S ribosomal protein L29	CKS_1580	*rpmC*	0	1
DSJ_00480	50S ribosomal protein L16	CKS_1581	*rplP*	0	2
DSJ_00485	30S ribosomal protein S3	CKS_1582	*rpsC*	0	1
DSJ_00490	**50S ribosomal protein L22**	**CKS_1583**	***rplV***	**0**	**0**
DSJ_00495	30S ribosomal protein S19	CKS_1584		0	1
DSJ_00500	50S ribosomal protein L2	CKS_1585	*rplB*	2	1
DSJ_00505	50S ribosomal protein L23	CKS_1586	*rplW*	0	1
DSJ_00510	**50S ribosomal protein L4**	**CKS_1587**	***rplD***	**0**	**0**

DSJ_00515	50S ribosomal protein L3	CKS_1588	*rplC*	2	1
DSJ_00520	**30S ribosomal protein S10**	**CKS_1589**	***rpsJ***	**0**	**0**
DSJ_00540	**translation elongation factor Tu**	**CKS_1593**		**0**	**0**
DSJ_00545	translation elongation factor G	CKS_1594	*fusA*	2	4
DSJ_00550	30S ribosomal protein S7	CKS_1595	*rpsG*	1	0
DSJ_00555	30S ribosomal protein S12	CKS_1596	*rpsL*	1	1
DSJ_00560	sulfurtransferase TusB	CKS_1597	*yheL*	1	3
DSJ_00565	sulfurtransferase TusC	CKS_1598	*yheM*	3	2
DSJ_00570	**sulfurtransferase TusD**	**CKS_1599**	***yheN***	**0**	**0**
DSJ_00650	hypothetical protein	CKS_1615	*yheU*	0.5	1.5
DSJ_00675	**resolvase**			**0**	**0**
DSJ_00680	**DEAD/DEAH box helicase**			**0**	**0**
DSJ_00685	**hypothetical protein**			**0**	**0**
DSJ_00715	**hypothetical protein**	**CKS_0841**		**0**	**0**
DSJ_00720	**antitoxin**			**0**	**0**
DSJ_00725	**plasmid maintenance protein CcdB**	**CKS_0839**		**0**	**0**
DSJ_00760	MFS transporter			4	4
DSJ_00770	**hypothetical protein**			**0**	**0**
DSJ_00840	hypothetical protein	CKS_0818		1	1
DSJ_00915	tryptophan--tRNA ligase	CKS_0854	*trpS*	0	1
DSJ_00930	DNA adenine methylase	CKS_0857	*dam*	5	4
DSJ_00935	cell division protein DamX	CKS_0858		1	1
DSJ_00940	3-dehydroquinate synthase	CKS_0859	*aroB*	1	0
DSJ_00945	shikimate kinase I	CKS_0860	*aroK*	1	1
DSJ_00960	hypothetical protein	CKS_5849		1	1
DSJ_01025	hypothetical protein	CKS_0877		6	5
DSJ_01120	aspartate-semialdehyde dehydrogenase	CKS_0900	*asd*	1	2
DSJ_01230	RNA polymerase factor sigma-32	CKS_0925	*rpoH*	0	1
DSJ_01275	**sulfurtransferase TusA**	**CKS_0934**	***sirA***	**0**	**0**
DSJ_01390	**hypothetical protein**			**0**	**0**
DSJ_01495	**hypothetical protein**			**0**	**0**
DSJ_01500	**hypothetical protein**			**0**	**0**
DSJ_01505	**hypothetical protein**			**0**	**0**
DSJ_01515	**hypothetical protein**			**0**	**0**
DSJ_01520	**hypothetical protein**			**0**	**0**
DSJ_01525	**hypothetical protein**			**0**	**0**
DSJ_01550	**hypothetical protein**			**0**	**0**
DSJ_01775	hypothetical protein	CKS_1038		0	1
DSJ_01900	uroporphyrinogen decarboxylase	CKS_0007	*hemE*	5.5	5
DSJ_01940	DNA-directed RNA polymerase subunit beta	CKS_0015	*rpoB*	1	4

DSJ_01945	50S ribosomal protein L7/L12	CKS_0016	*rplL*	0	1
DSJ_01950	**50S ribosomal protein L10**	**CKS_0017**	***rplJ***	**0**	**0**
DSJ_01955	50S ribosomal protein L1	CKS_0018	*rplA*	0	1
DSJ_01960	**50S ribosomal protein L11**	**CKS_0019**	***rplK***	**0**	**0**
DSJ_01965	nusG	CKS_0020	*nusG*	2	1
DSJ_01970	preprotein translocase subunit SecE	CKS_0021	*secE*	2	9
DSJ_01975	**translation elongation factor Tu**	**CKS_0022**		**0**	**0**
DSJ_02005	biotin--[acetyl-CoA-carboxylase] synthetase	CKS_0029	*birA*	2	2
DSJ_02010	UDP-N-acetylenolpyruvoylglucosamine reductase	CKS_0030	*murB*	3	1
DSJ_02130	tRNA uridine-5-carboxymethylaminomethyl(34) synthesis enzyme MnmG	CKS_1215		4	2
DSJ_02140	ATP F0F1 synthase subunit I	CKS_1213	*atpI*	0	1
DSJ_02145	F0F1 ATP synthase subunit A	CKS_1212	*atpB*	0	3
DSJ_02150	ATP F0F1 synthase subunit C	CKS_1211	*atpE*	0	1
DSJ_02155	F0F1 ATP synthase subunit B	CKS_1210	*atpF*	1	3
DSJ_02160	F0F1 ATP synthase subunit delta	CKS_1209	*atpH*	2	1
DSJ_02165	F0F1 ATP synthase subunit alpha	CKS_1208	*atpA*	3	1
DSJ_02170	F0F1 ATP synthase subunit gamma	CKS_1207	*atpG*	1	0
DSJ_02175	F0F1 ATP synthase subunit beta	CKS_1206	*atpD*	1	3
DSJ_02185	UDP-N-acetylglucosamine diphosphorylase/glucosamine-1-phosphate N-acetyltransferase	CKS_1204	*glmU*	3	3
DSJ_02220	phosphate transport system regulatory protein PhoU	CKS_1196	*phoU*	2	0
DSJ_02335	**hypothetical protein**			**0**	**0**
DSJ_02365	**hypothetical protein**	**CKS_1160**		**0**	**0**
DSJ_02380	**hypothetical protein**			**0**	**0**
DSJ_02420	deoxyuridine 5'-triphosphate nucleotidohydrolase	CKS_1150	*dut*	1	6
DSJ_02425	bifunctional phosphopantothenoylcysteine decarboxylase/phosphopantothenate synthase	CKS_1149	*dfp*	4	6
DSJ_02450	pantetheine-phosphate adenylyltransferase	CKS_1143	*coaD*	1	0
DSJ_02460	3-deoxy-D-manno-octulosonic acid transferase	CKS_1141	*waaA*	1	3
DSJ_02495	**hypothetical protein**			**0**	**0**
DSJ_02535	ADP-glyceromanno-heptose 6-epimerase	CKS_1120	*rfaD*	1	9
DSJ_02540	ADP-heptose--LPS heptosyltransferase	CKS_1119	*rfaF*	5	5

DSJ_02580	protein-export chaperone SecB	CKS_1111	*secB*	0	2
DSJ_02720	50S ribosomal protein L31	CKS_1080		1	0
DSJ_02795	DNA-binding transcriptional regulator FabR	CKS_1059	*fabR*	7	9
DSJ_02815	**glutamate racemase**	**CKS_1054**	***murI***	**0**	**0**
DSJ_03055	**chorismate lyase**	**CKS_4681**	***ubiC***	**0**	**0**
DSJ_03185	**hypothetical protein**			**0**	**0**
DSJ_03670	VOC family protein			1	7
DSJ_03790	**groES**	**CKS_4848**	***groS***	**0**	**0**
DSJ_03810	elongation factor P	CKS_4852	*efp*	1	2
DSJ_03830	elongation factor P lysine(34) lysyltransferase	CKS_4856	*poxA*	1	4
DSJ_03835	**outer membrane-stress sensor serine endopeptidase DegS**	**CKS_4857**	***degS***	**0**	**0**
DSJ_03845	hypothetical protein	CKS_4859	*yhcB*	2	7
DSJ_03855	**50S ribosomal protein L13**	**CKS_4861**	***rplM***	**0**	**0**
DSJ_03860	30S ribosomal protein S9	CKS_4862	*rpsI*	0	1
DSJ_03895	phosphohistidinoprotein-hexose phosphotransferase	CKS_4870	*npr*	0.5	0
DSJ_03920	LPS export ABC transporter ATP-binding protein	CKS_4875		1	1
DSJ_03925	lipopolysaccharide ABC transporter substrate-binding protein LptA	CKS_4876		1	4
DSJ_03930	**LPS export ABC transporter periplasmic protein LptC**	**CKS_4877**		**0**	**0**
DSJ_03935	3-deoxy-D-manno-octulosonate 8-phosphate phosphatase	CKS_4878	*kdsC*	6	0
DSJ_03940	D-arabinose 5-phosphate isomerase	CKS_4879	*kdsD*	3	4
DSJ_03970	**anti-sigma B factor antagonist**	**CKS_4885**		**0**	**0**
DSJ_03975	hypothetical protein	CKS_4886	*yrbA*	1	1
DSJ_03980	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	CKS_4887	*murA*	4	7
DSJ_03985	**DNA-binding protein**	**CKS_4888**	***sfsB***	**0**	**0**
DSJ_03990	octaprenyl diphosphate synthase	CKS_4889	*ispB*	2	1
DSJ_03995	50S ribosomal protein L21	CKS_4890	*rplU*	2	1
DSJ_04000	**50S ribosomal protein L27**	**CKS_4891**	***rpmA***	**0**	**0**
DSJ_04005	obgE	CKS_4892	*obgE*	1	2
DSJ_04025	23S rRNA (uridine(2552)-2'-O)-methyltransferase	CKS_4897	*rlmE*	4	7
DSJ_04030	hflB	CKS_4898	*ftsH*	7	4
DSJ_04035	dihydropteroate synthase	CKS_4899	*folP*	1	1
DSJ_04040	phosphoglucosamine mutase	CKS_4900	*glmM*	3	1
DSJ_04045	**preprotein translocase subunit SecG**	**CKS_4901**	***secG***	**0**	**0**

DSJ_04060	**ribosome maturation factor RimP**	**CKS_4904**	***rimP***	**0**	**0**
DSJ_04065	nusA	CKS_4905	*nusA*	0	2
DSJ_04070	translation initiation factor IF-2	CKS_4906	*infB*	1	1
DSJ_04075	ribosome-binding factor A	CKS_4907	*rbfA*	1	1
DSJ_04250	valine--tRNA ligase	CKS_4946	*valS*	1	5
DSJ_04255	DNA polymerase III subunit chi	CKS_4947	*holC*	0	2
DSJ_04265	lipopolysaccharide ABC transporter permease LptF	CKS_4949		0	4
DSJ_04270	lipopolysaccharide ABC transporter permease LptG	CKS_4950		1	2
DSJ_04445	**DNA polymerase III subunit psi**	**CKS_4990**		**0**	**0**
DSJ_04595	PTS mannitol transporter subunit IIB	CKS_1424		4	9
DSJ_04635	**Holliday junction DNA helicase RuvA**	**CKS_1415**	***yqgF***	**0**	**0**
DSJ_04640	DUF179 domain-containing protein	CKS_1414	*yqgE*	0	1
DSJ_04675	methionine adenosyltransferase	CKS_1407	*metK*	0	1
DSJ_04705	phosphoglycerate kinase	CKS_1401	*pgk*	0	5
DSJ_04710	class II fructose-bisphosphate aldolase	CKS_1400	*fbaA*	0	2
DSJ_04730	**hypothetical protein**			**0**	**0**
DSJ_04775	2-octaprenyl-6-methoxyphenyl hydroxylase	CKS_1388	*ubiH*	1	2
DSJ_04780	FAD-dependent 2-octaprenylphenol hydroxylase	CKS_1387	*visC*	4	7
DSJ_04820	tRNA-modifying protein YgfZ	CKS_1379	*ygfZ*	0	1
DSJ_04930	prolipoprotein diacylglyceryl transferase	CKS_1354	*lgt*	2	0
DSJ_04935	thymidylate synthase	CKS_1353	*thyA*	1	2
DSJ_04960	exodeoxyribonuclease V subunit gamma	CKS_1347	*recC*	1	7
DSJ_04965	pitrilysin	CKS_1345	*ptrA*	2	0
DSJ_04970	exodeoxyribonuclease V subunit beta	CKS_1344	*recB*	4	4
DSJ_05065	hypothetical protein	CKS_1325		1	6
DSJ_05070	tRNA pseudouridine(65) synthase TruC	CKS_1324	*truC*	2	7
DSJ_05130	CTP synthase	CKS_1311	*pyrG*	2	2
DSJ_05135	phosphopyruvate hydratase	CKS_1310	*eno*	9	7
DSJ_05275	**cell division protein FtsB**	**CKS_1279**	***ftsB***	**0**	**0**
DSJ_05280	**2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase**	**CKS_1278**	***ispD***	**0**	**0**
DSJ_05285	**2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase**	**CKS_1277**	***ispF***	**0**	**0**
DSJ_05360	alanine--tRNA ligase	CKS_1260	*alaS*	0	2
DSJ_05365	**carbon storage regulator**	**CKS_1259**		**0**	**0**
DSJ_05420	signal recognition particle protein	CKS_1248	*ffh*	3	7
DSJ_05425	**30S ribosomal protein S16**	**CKS_1247**	***rpsP***	**0**	**0**
DSJ_05430	rimM	CKS_1246	*rimM*	2	0

DSJ_05435	tRNA (guanosine(37)-N1)-methyltransferase TrmD	CKS_1245	*trmD*	2	0
DSJ_05440	50S ribosomal protein L19	CKS_1244	*rplS*	4	1
DSJ_05540	**hypothetical protein**	**CKS_5018**		**0**	**0**
DSJ_05580	**hypothetical protein**			**0**	**0**
DSJ_05585	**hypothetical protein**			**0**	**0**
DSJ_05595	hypothetical protein	CKS_5028		3	3
DSJ_05715	**hypothetical protein**	**CKS_5052**	***yaiA***	**0**	**0**
DSJ_05735	pyrroline-5-carboxylate reductase	CKS_5056	*proC*	5	6
DSJ_05795	peptide chain release factor 3	CKS_5069	*prfC*	1	4
DSJ_05810	**hypothetical protein**	**CKS_5072**		**0**	**0**
DSJ_05840	**hypothetical protein**			**0**	**0**
DSJ_05845	**group II intron reverse transcriptase/maturase**	**CKS_5081**		**0**	**0**
DSJ_06005	nhaA	CKS_5118	*nhaA*	2	1
DSJ_06015	riboflavin biosynthesis protein RibF	CKS_5121	*ribF*	1	0
DSJ_06020	isoleucine--tRNA ligase	CKS_5122	*ileS*	0	5
DSJ_06025	**signal peptidase II**	**CKS_5123**	***lspA***	**0**	**0**
DSJ_06030	**peptidylprolyl isomerase**	**CKS_5124**	***fkpB***	**0**	**0**
DSJ_06040	4-hydroxy-tetrahydrodipicolinate reductase	CKS_5126	*dapB*	0	1
DSJ_06060	**folA**	**CKS_5131**	***folA***	**0**	**0**
DSJ_06200	cell division/cell wall cluster transcriptional repressor MraZ	CKS_5161	*mraZ*	0	0.5
DSJ_06205	16S rRNA (cytosine(1402)-N(4))-methyltransferase	CKS_5162		0	1.5
DSJ_06210	**cell division protein FtsL**	**CKS_5163**	***ftsL***	**0**	**0**
DSJ_06215	peptidoglycan glycosyltransferase FtsI	CKS_5164	*ftsI*	0	5
DSJ_06220	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2, 6-diaminopimelate ligase	CKS_5165	*murE*	0	4
DSJ_06225	UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D- alanine ligase	CKS_5166	*murF*	5	1
DSJ_06230	phospho-N-acetylmuramoyl-pentapeptide- transferase	CKS_5167	*mraY*	1	4.5
DSJ_06235	murD	CKS_5168	*murD*	0	4.5
DSJ_06240	cell division protein FtsW	CKS_5169	*ftsW*	2	0
DSJ_06245	undecaprenyldiphospho-muramoylpentapeptide beta-N-acetylglucosaminyltransferase	CKS_5170	*murG*	1	1
DSJ_06250	UDP-N-acetylmuramate--L-alanine ligase	CKS_5171	*murC*	2	1
DSJ_06255	ddl	CKS_5172	*ddlB*	2	6
DSJ_06260	cell division protein FtsQ	CKS_5173	*ftsQ*	0	1

DSJ_06265	cell division protein FtsA	CKS_5174	*ftsA*	1	5
DSJ_06270	cell division protein FtsZ	CKS_5175	*ftsZ*	4	9
DSJ_06275	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosaminedeacetylase	CKS_5176	*lpxC*	1	0
DSJ_06280	**SecA regulator SecM**	**CKS_5177**		**0**	**0**
DSJ_06295	**DNA gyrase inhibitor YacG**	**CKS_5180**	***yacG***	**0**	**0**
DSJ_06305	dephospho-CoA kinase	CKS_5182	*coaE*	0	1
DSJ_06370	pyruvate dehydrogenase (acetyl-transferring), homodimeric type	CKS_5196	*aceE*	6	8
DSJ_06380	dihydrolipoyl dehydrogenase	CKS_5198	*lpd*	3	5
DSJ_06515	2-amino-4-hydroxy-6-hydroxymethyldihydropteridine diphosphokinase	CKS_5232	*folK*	2	3
DSJ_06520	polynucleotide adenylyltransferase	CKS_5233	*pcnB*	0	3
DSJ_06530	RNA polymerase-binding protein DksA	CKS_5235	*dksA*	6	1
DSJ_06560	iron-hydroxamate transporter ATP-binding subunit	CKS_5241	*fhuC*	0	3
DSJ_06565	iron-hydroxamate transporter substrate-binding subunit	CKS_5242		0	1
DSJ_06570	Fe3+-hydroxamate ABC transporter permease FhuB	CKS_5243	*fhuB*	1	6
DSJ_06575	glutamate-1-semialdehyde-2,1-aminomutase	CKS_5244	*hemL*	1	5
DSJ_06580	iron-sulfur cluster insertion protein ErpA	CKS_5245		1	1
DSJ_06625	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase	CKS_5255	*dapD*	0	5
DSJ_06640	30S ribosomal protein S2	CKS_5258	*rpsB*	1	0
DSJ_06645	translation elongation factor Ts	CKS_5259	*tsf*	2	0
DSJ_06650	UMP kinase	CKS_5260	*pyrH*	0	1
DSJ_06655	ribosome recycling factor	CKS_5261	*frr*	0	2
DSJ_06660	1-deoxy-D-xylulose-5-phosphate reductoisomerase	CKS_5262	*dxr*	5	9
DSJ_06665	(2E,6E)-farnesyl- diphosphate-specific ditrans,polycis-undecaprenyl-diphosphate synthase	CKS_5263	*ispU*	0	1
DSJ_06670	phosphatidate cytidylyltransferase	CKS_5264	*cdsA*	0	2
DSJ_06675	**RIP metalloprotease RseP**	**CKS_5265**	***rseP***	**0**	**0**
DSJ_06685	molecular chaperone	CKS_5267		1	1
DSJ_06690	UDP-3-O-(3-hydroxymyristoyl)glucosamine N-acyltransferase	CKS_5268	*lpxD*	1	0
DSJ_06695	3-hydroxyacyl-[acyl-carrier-protein] dehydrataseFabZ	CKS_5269	*fabZ*	0	1
DSJ_06700	acyl-[acyl-carrier-protein]--UDP-N-	CKS_5270	*lpxA*	0	2

	acetylglucosamine O-acyltransferase				
DSJ_06705	**lipid-A-disaccharide synthase**	**CKS_5271**	***lpxB***	**0**	**0**
DSJ_06715	DNA polymerase III subunit alpha	CKS_5273	*dnaE*	5	6
DSJ_06735	tRNA lysidine(34) synthetase TilS	CKS_5277		1	2
DSJ_06760	proline--tRNA ligase	CKS_5282	*proS*	0	3
DSJ_06845	pssA	CKS_1446	*pssA*	1	4
DSJ_06915	NrdH-redoxin	CKS_1464	*nrdH*	0	1
DSJ_07040	ubiquinone-binding protein	CKS_1677	*ratA*	1	0
DSJ_07060	**NAD(+) kinase**	**CKS_1673**	***nadK***	**0**	**0**
DSJ_07095	**RNA polymerase sigma factor RpoE**	**CKS_1665**	***rpoE***	**0**	**0**
DSJ_07100	anti-sigma factor	CKS_1664	*rseA*	4.5	9
DSJ_07115	elongation factor 4	CKS_1661	*lepA*	4	5
DSJ_07120	S26 family signal peptidase	CKS_1660	*lepB*	2	2
DSJ_07125	ribonuclease III	CKS_1659	*rnc*	2	2
DSJ_07130	GTPase Era	CKS_1658	*era*	1	1
DSJ_07140	pyridoxine 5'-phosphate synthase	CKS_1656	*pdxJ*	0	1
DSJ_07145	acpS	CKS_1655	*acpS*	3	2
DSJ_07230	hypothetical protein	CKS_1688		5	2.5
DSJ_07370	hypothetical protein	CKS_1496		2	0
DSJ_07390	tRNA adenosine(34) deaminase TadA	CKS_1500	*tadA*	0	2
DSJ_07445	serine hydroxymethyltransferase	CKS_1510	*glyA*	4	4
DSJ_07520	nucleoside-diphosphate kinase	CKS_1527	*ndk*	2	2
DSJ_07540	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase	CKS_1531	*ispG*	1	1
DSJ_07545	histidine--tRNA ligase	CKS_1532	*hisS*	1	2
DSJ_07590	glutamine-hydrolyzing GMP synthase	CKS_1542	*guaA*	1	1
DSJ_07600	hypothetical protein			4	9
DSJ_07655	**ArsC family reductase**	**CKS_1738**	***yffB***	**0**	**0**
DSJ_07660	succinyl-diaminopimelate desuccinylase	CKS_1737	*dapE*	1	1
DSJ_07695	4-hydroxy-tetrahydrodipicolinate synthase	CKS_1730	*dapA*	0	1
DSJ_07770	**hypothetical protein**			**0**	**0**
DSJ_07785	DnaA regulatory inactivator Hda	CKS_1710		2	5
DSJ_07985	lipoyl synthase	CKS_3723	*lipA*	3	1
DSJ_07995	hypothetical protein	CKS_3725	*ybeD*	5	8
DSJ_08010	rod shape-determining protein RodA	CKS_3728	*mrdB*	4	3
DSJ_08020	**23S rRNA (pseudouridine(1915)-N(3))-methyltransferase RlmH**	**CKS_3730**		**0**	**0**
DSJ_08025	**ribosome silencing factor**	**CKS_3731**		**0**	**0**
DSJ_08035	DNA polymerase III subunit delta	CKS_3733	*holA*	1	0
DSJ_08040	LPS assembly lipoprotein LptE	CKS_3734		0	1
DSJ_08045	leucine--tRNA ligase	CKS_3735	*leuS*	4	2

DSJ_08090	apolipoprotein N-acyltransferase	CKS_3823	*lnt*	2	0
DSJ_08095	**magnesium/cobalt efflux protein**	**CKS_3824**	***ybeX***	**0**	**0**
DSJ_08100	**rRNA maturation RNase YbeY**	**CKS_3825**	***ybeY***	**0**	**0**
DSJ_08105	nucleoside triphosphate hydrolase	CKS_3826	*ybeZ*	4	4
DSJ_08115	**ubiF**	**CKS_3829**	***ubiF***	**0**	**0**
DSJ_08195	flavodoxin	CKS_3761	*fldA*	2	0
DSJ_08215	phosphoglucomutase, alpha-D-glucose phosphate-specific	CKS_3765	*pgm*	1	5
DSJ_08300	succinate dehydrogenase cytochrome b556 large subunit	CKS_3784		2	3
DSJ_08305	succinate dehydrogenase, hydrophobic membrane anchor protein	CKS_3785	*sdhD*	7	8
DSJ_08315	succinate dehydrogenase iron-sulfur subunit	CKS_3787	*sdhB*	5	5
DSJ_08325	dihydrolipoamide succinyltransferase	CKS_3790	*sucB*	2	1
DSJ_08360	tol-pal system-associated acyl-CoA thioesterase	CKS_3797	*ybgC*	1	1
DSJ_08365	protein TolQ	CKS_3798	*tolQ*	1	0
DSJ_08370	protein TolR	CKS_3799	*tolR*	1	0
DSJ_08380	Tol-Pal system beta propeller repeat protein TolB	CKS_3801	*tolB*	0	2
DSJ_08580	molybdenum cofactor biosynthesis protein B	CKS_3859	*moaB*	2	1
DSJ_08815	phage virion morphogenesis protein	CKS_5687		6	9.5
DSJ_08825	**hypothetical protein**	**CKS_5685**		**0**	**0**
DSJ_08845	hypothetical protein	CKS_5681		5	7
DSJ_08865	hypothetical protein	CKS_3912		0	1
DSJ_09020	DNA-binding protein	CKS_3952		3	5
DSJ_09175	**hypothetical protein**			**0**	**0**
DSJ_09300	hypothetical protein			0	1
DSJ_09350	putrescine ABC transporter permease PotI			9	5
DSJ_09445	translation initiation factor IF-1	CKS_3971	*infA*	0	2
DSJ_09480	**lolA**	**CKS_3979**	***lolA***	**0**	**0**
DSJ_09490	serine--tRNA ligase	CKS_3981	*serS*	2	7
DSJ_09535	phosphoserine transaminase	CKS_3990	*serC*	0	3
DSJ_09540	3-phosphoshikimate 1-carboxyvinyltransferase	CKS_3991	*aroA*	4	3
DSJ_09545	cytidylate kinase	CKS_3992	*cmk*	1	4
DSJ_09550	30S ribosomal protein S1	CKS_3993	*rpsA*	2	2
DSJ_09565	lipid ABC transporter permease/ATP-binding protein	CKS_3998	*msbA*	7	2
DSJ_09580	**hypothetical protein**	**CKS_4001**	***ycaR***	**0**	**0**

DSJ_09585	3-deoxy-manno-octulosonate cytidylyltransferase	CKS_4002	*kdsB*	1	2
DSJ_09595	tRNA uridine 5-oxyacetic acid(34) methyltransferase CmoM	CKS_4004	*smtA*	3	9
DSJ_09600	condensin subunit MukF	CKS_4005	*mukF*	3	3
DSJ_09605	chromosome partitioning protein MukE	CKS_4006	*mukE*	2	5
DSJ_09640	**asparagine--tRNA ligase**	**CKS_4013**	***asnS***	**0**	**0**
DSJ_09730	**3-hydroxyacyl-[acyl-carrier-protein] dehydrataseFabA**	**CKS_4033**	***fabA***	**0**	**0**
DSJ_09805	**sulfurtransferase TusE**	**CKS_4049**	***yccK***	**0**	**0**
DSJ_09890	**sodium:solute symporter**	**CKS_4066**		**0**	**0**
DSJ_09920	hypothetical protein	CKS_4071		1	0
DSJ_09925	hypothetical protein	CKS_4072		1	0
DSJ_09940	**hypothetical protein**	**CKS_4075**		**0**	**0**
DSJ_09950	repressor			8	2
DSJ_09955	**hypothetical protein**	**CKS_4076**		**0**	**0**
DSJ_10035	**hypothetical protein**	**CKS_4095**		**0**	**0**
DSJ_10050	**hypothetical protein**	**CKS_0979**		**0**	**0**
DSJ_10055	**hypothetical protein**			**0**	**0**
DSJ_10060	**hypothetical protein**			**0**	**0**
DSJ_10075	**hypothetical protein**	**CKS_0981**		**0**	**0**
DSJ_10080	**hypothetical protein**	**CKS_0983**		**0**	**0**
DSJ_10085	**hypothetical protein**	**CKS_0984**		**0**	**0**
DSJ_10115	**hypothetical protein**			**0**	**0**
DSJ_10165	head-tail adaptor	CKS_4118		3	4
DSJ_10330	**CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase**	**CKS_4152**	***pgsA***	**0**	**0**
DSJ_10365	hypothetical protein	CKS_4160		1	2
DSJ_10390	hypothetical protein			8	3
DSJ_10460	lipid A biosynthesis lauroyl acyltransferase	CKS_4184	*lpxL*	4	3
DSJ_10495	**DNA damage-inducible protein I**	**CKS_4191**	***dinI***	**0**	**0**
DSJ_10525	murein biosynthesis integral membrane protein MurJ	CKS_4198		1	6
DSJ_10640	hypothetical protein	CKS_4225	*yceD*	1	2
DSJ_10645	50S ribosomal protein L32	CKS_4226	*rpmF*	0	1
DSJ_10650	phosphate acyltransferase	CKS_4227	*plsX*	0	1
DSJ_10655	3-oxoacyl-ACP synthase	CKS_4228	*fabH*	1	0
DSJ_10660	**[acyl-carrier-protein] S-malonyltransferase**	**CKS_4229**	***fabD***	**0**	**0**
DSJ_10670	**acyl carrier protein**	**CKS_4231**	***acpP***	**0**	**0**
DSJ_10690	dTMP kinase	CKS_4236	*tmk*	2	1
DSJ_10695	**DNA polymerase III subunit delta'**	**CKS_4237**	***holB***	**0**	**0**

DSJ_10765	outer membrane-specific lipoprotein transporter subunit LolC	CKS_4253	*lolC*	2	0
DSJ_10770	lolD	CKS_4254	*lolD*	0	1
DSJ_10775	lipoprotein transporter subunit LolE	CKS_4255	*lolE*	1	2
DSJ_10805	adenylosuccinate lyase	CKS_4263	*purB*	0	2
DSJ_10810	lysogenization regulator HflD	CKS_4264	*hflD*	0	3
DSJ_10815	tRNA 2-thiouridine(34) synthase MnmA	CKS_4265	*mnmA*	4	2
DSJ_10900	**hypothetical protein**	**CKS_4281**		**0**	**0**
DSJ_10915	transcriptional regulator	CKS_1641		3	1
DSJ_10970	hypothetical protein	CKS_2362		0	1
DSJ_11340	methyltransferase	CKS_2441		1	1
DSJ_11350	**hypothetical protein**	**CKS_2443**		**0**	**0**
DSJ_11780	**hypothetical protein**	**CKS_2540**		**0**	**0**
DSJ_11810	ferrochelatase	CKS_2547	*hemH*	3	0
DSJ_12000	**hypothetical protein**	**CKS_5716**		**0**	**0**
DSJ_12005	**hypothetical protein**	**CKS_2589**		**0**	**0**
DSJ_12015	**hypothetical protein**	**CKS_2591**		**0**	**0**
DSJ_12080	hypothetical protein			9	8
DSJ_12255	hypothetical protein	CKS_2634		0	4
DSJ_12345	**phage repressor protein**	**CKS_2648**		**0**	**0**
DSJ_12555	ribose-phosphate pyrophosphokinase	CKS_3321	*prs*	3	2
DSJ_12560	4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase	CKS_3322	*ispE*	0	2
DSJ_12565	lipoprotein localization factor LolB	CKS_3323	*lolB*	4	3
DSJ_12570	glutamyl-tRNA reductase	CKS_3325	*hemA*	0	2
DSJ_12575	peptide chain release factor 1	CKS_3326	*prfA*	2	1
DSJ_12580	protein-(glutamine-N5) methyltransferase, release factor-specific	CKS_3327	*prmC*	2	9
DSJ_12595	3-deoxy-8-phosphooctulonate synthase	CKS_3330	*kdsA*	1	1
DSJ_12650	hypothetical protein			0	1
DSJ_12795	NAD(+) synthase	CKS_3377	*nadE*	1	1
DSJ_12850	threonine--tRNA ligase	CKS_3388	*thrS*	1	4
DSJ_12855	**translation initiation factor IF-3**	**CKS_3389**	***infC***	**0**	**0**
DSJ_12860	**50S ribosomal protein L35**	**CKS_3390**	***rpmI***	**0**	**0**
DSJ_12865	50S ribosomal protein L20	CKS_3391	*rplT*	0	2
DSJ_12875	phenylalanine--tRNA ligase subunit alpha	CKS_3393	*pheS*	0	1
DSJ_12880	phenylalanine--tRNA ligase subunit beta	CKS_3394	*pheT*	1	8
DSJ_12885	**integration host factor subunit alpha**	**CKS_3395**	***ihfA***	**0**	**0**
DSJ_12960	Fe-S cluster assembly scaffold SufA	CKS_3412	*sufA*	0	1
DSJ_12965	Fe-S cluster assembly protein SufB	CKS_3413	*sufB*	2	0
DSJ_12970	Fe-S cluster assembly ATPase SufC	CKS_3414	*sufC*	1	1

DSJ_12975	**FeS cluster assembly protein SufD**	**CKS_3415**	***sufD***	**0**	**0**
DSJ_12980	bifunctional cysteine desulfurase/selenocysteinelyase	CKS_3416	*sufS*	1	0
DSJ_13025	riboflavin synthase subunit alpha	CKS_3426	*ribC*	1	0
DSJ_13055	**monothiol glutaredoxin, Grx4 family**	**CKS_3437**	***grxD***	**0**	**0**
DSJ_13110	**hypothetical protein**	**CKS_2123**		**0**	**0**
DSJ_13145	**hypothetical protein**	**CKS_3456**	***ydhI***	**0**	**0**
DSJ_13240	pyridoxamine 5'-phosphate oxidase	CKS_3477	*pdxH*	9	5
DSJ_13245	tyrosine--tRNA ligase	CKS_3478	*tyrS*	0	1
DSJ_13275	electron transport complex subunit RsxE	CKS_3485	*rsxE*	1	4
DSJ_13300	electron transport complex subunit RsxA	CKS_3490	*rsxA*	7	7
DSJ_13335	hypothetical protein	CKS_3498	*yaiL*	1	1
DSJ_13425	hypothetical protein	CKS_3523		1	3
DSJ_13545	DNA replication terminus site-binding protein	CKS_3554	*tus*	1	3
DSJ_13570	transcriptional regulator	CKS_3562		2	5
DSJ_13725	hypothetical protein	CKS_3600		1	4.5
DSJ_13775	**LD-carboxypeptidase**			**0**	**0**
DSJ_13780	hypothetical protein	CKS_2950		5	8
DSJ_13950	**hypothetical protein**	**CKS_2911**		**0**	**0**
DSJ_14125	hypothetical protein	CKS_2874	*ybiI*	6	5
DSJ_14230	serine protease	CKS_2846	*ydgD*	9	5
DSJ_14395	enoyl-[acyl-carrier-protein] reductase	CKS_2811	*fabI*	1	2
DSJ_14430	**hypothetical protein**	**CKS_2803**		**0**	**0**
DSJ_14470	lipopolysaccharide assembly protein LapB	CKS_2796	*yciM*	1	1
DSJ_14475	**hypothetical protein**	**CKS_2795**	***yciS***	**0**	**0**
DSJ_14485	GTP cyclohydrolase II	CKS_2793	*ribA*	1	0
DSJ_14495	**hypothetical protein**	**CKS_2791**		**0**	**0**
DSJ_14505	DNA topoisomerase I	CKS_2789	*topA*	4	4
DSJ_14625	**TonB system transport protein TonB**	**CKS_2761**		**0**	**0**
DSJ_14645	dsDNA-mimic protein	CKS_2755	*yciU*	2	0
DSJ_14725	UTP--glucose-1-phosphate uridylyltransferase	CKS_2737	*galU*	2	4
DSJ_14805	**type I glyceraldehyde-3-phosphate dehydrogenase**	**CKS_2720**	***gapA***	**0**	**0**
DSJ_14865	alanine racemase	CKS_2706	*dadX*	1	5.5
DSJ_14915	**cell division topological specificity factor MinE**	**CKS_2695**		**0**	**0**
DSJ_15050	MFS transporter			5	4
DSJ_15275	Holliday junction DNA helicase RuvB	CKS_1822	*ruvB*	1	5
DSJ_15280	Holliday junction DNA helicase RuvA	CKS_1823	*ruvA*	1	0

DSJ_15285	crossover junction endodeoxyribonuclease RuvC	CKS_1824	*ruvC*	3	8
DSJ_15290	hypothetical protein	CKS_1825	*yebC*	1	3
DSJ_15295	**hypothetical protein**	**CKS_1826**		**0**	**0**
DSJ_15300	dihydroneopterin triphosphate diphosphatase	CKS_1827	*nudB*	1	0
DSJ_15305	**aspartate--tRNA ligase**	**CKS_1828**	***aspS***	**0**	**0**
DSJ_15555	hypothetical protein			0	4
DSJ_15690	**host nuclease inhibitor protein**	**CKS_2333**		**0**	**0**
DSJ_15790	**hypothetical protein**			**0**	**0**
DSJ_15945	hypothetical protein	CKS_2273		1	1
DSJ_15955	glycosyl transferase family 2	CKS_2271		1	0
DSJ_15960	hypothetical protein	CKS_2270		1	2
DSJ_15965	hypothetical protein	CKS_2269		4	2
DSJ_15985	glucose-1-phosphate thymidylyltransferase	CKS_2266		4	4
DSJ_15990	dTDP-glucose 4,6-dehydratase	CKS_2265	*rfbB*	1	4
DSJ_16080	hypothetical protein			3	1
DSJ_16255	peroxiredoxin	CKS_2199	*ahpC*	4	8
DSJ_16330	GTP cyclohydrolase I FolE	CKS_2180	*folE*	2	3
DSJ_16365	**group II intron reverse transcriptase/maturase**			**0**	**0**
DSJ_16370	**hypothetical protein**			**0**	**0**
DSJ_16465	hypothetical protein	CKS_2150		6	9
DSJ_16485	50S ribosomal protein L25	CKS_2146	*rplY*	1	0
DSJ_16495	hypothetical protein	CKS_2144	*yejL*	0	1
DSJ_16585	head-tail adaptor			3	1
DSJ_16590	hypothetical protein			0	4
DSJ_16620	**hypothetical protein**			**0**	**0**
DSJ_16625	**HNH endonuclease**			**0**	**0**
DSJ_16630	**hypothetical protein**			**0**	**0**
DSJ_16635	**glycosyltransferase**			**0**	**0**
DSJ_16640	**hypothetical protein**			**0**	**0**
DSJ_16645	**hypothetical protein**			**0**	**0**
DSJ_16715	repressor			0	4
DSJ_16745	hypothetical protein			1	3
DSJ_16760	hypothetical protein			6	9
DSJ_16775	hypothetical protein	CKS_2140		0	2
DSJ_16785	**transcriptional regulator**			**0**	**0**
DSJ_16965	ribonucleoside-diphosphate reductase subunit alpha	CKS_2099	*nrdA*	3	5
DSJ_16970	ribonucleotide-diphosphate reductase	CKS_2098	*nrdB*	1	1

	subunit beta				
DSJ_17030	NADH-quinone oxidoreductase subunit K	CKS_2085	*nuoK*	7	6
DSJ_17205	3-octaprenyl-4-hydroxybenzoate carboxy-lyase	CKS_2048	*ubiX*	0	1
DSJ_17225	**bifunctional tetrahydrofolate synthase/dihydrofolate synthase**	**CKS_2044**	***folC***	**0**	**0**
DSJ_17230	acetyl-CoA carboxylase subunit beta	CKS_2043	*accD*	2	1
DSJ_17300	beta-ketoacyl-[acyl-carrier-protein] synthase I	CKS_2027	*fabB*	8	8
DSJ_17330	**chorismate synthase**	**CKS_2021**	***aroC***	**0**	**0**
DSJ_17335	ribosomal protein L3 N(5)-glutamine methyltransferase	CKS_2020	*prmB*	2	1
DSJ_17815	**hypothetical protein**	**CKS_1923**		**0**	**0**
DSJ_17875	**hypothetical protein**	**CKS_1912**		**0**	**0**
DSJ_18005	translation initiation factor IF-2	CKS_3668		1	1
DSJ_18050	hypothetical protein	CKS_3656		1	0
DSJ_18270	glutamate--tRNA ligase	CKS_1769	*gltX*	2	2
DSJ_18320	**ligA**	**CKS_0044**	***ligA***	**0**	**0**
DSJ_18325	cell division protein ZipA	CKS_0045	*zipA*	1	1
DSJ_18410	coproporphyrinogen III oxidase	CKS_0064	*hemF*	2	2
DSJ_18445	head-tail adaptor	CKS_0074		1.5	8
DSJ_18480	**hypothetical protein**	**CKS_0080**		**0**	**0**
DSJ_18485	**HNH endonuclease**	**CKS_0081**		**0**	**0**
DSJ_18490	**hypothetical protein**	**CKS_0082**		**0**	**0**
DSJ_18495	**glycosyltransferase**	**CKS_0083**		**0**	**0**
DSJ_18500	**hypothetical protein**	**CKS_0084**		**0**	**0**
DSJ_18505	**hypothetical protein**	**CKS_0085**		**0**	**0**
DSJ_18510	Rz lytic protein	CKS_0086		4.5	9
DSJ_18550	**hypothetical protein**	**CKS_0093**		**0**	**0**
DSJ_18630	**sodium:solute symporter**	**CKS_0105**		**0**	**0**
DSJ_18670	bifunctional methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase	CKS_0113	*folD*	0	1
DSJ_18690	peptidylprolyl isomerase	CKS_0117	*ppiB*	1	1
DSJ_18695	UDP-2,3-diacylglucosamine diphosphatase	CKS_0118	*lpxH*	1	3
DSJ_18700	5-(carboxyamino)imidazole ribonucleotide mutase	CKS_0119	*purE*	5	3
DSJ_18805	prevent-host-death family protein	CKS_0142		2	4
DSJ_18980	adenylate kinase	CKS_0178	*adk*	2	1
DSJ_18995	**YbaB/EbfC family nucleoid-associated protein**	**CKS_0181**	***ybaB***	**0**	**0**

DSJ_19000	DNA polymerase III subunit gamma/tau	CKS_0182	*dnaX*	1	4
DSJ_19020	DUF2496 domain-containing protein	CKS_0186	*ybaM*	2	6
DSJ_19030	efflux transporter periplasmic adaptor subunit	CKS_0188	*acrA*	0	2
DSJ_19035	aminoglycoside/multidrug transporter permease	CKS_0189		9	2
DSJ_19165	ATP-dependent protease ATP-binding subunit ClpX	CKS_0218	*clpX*	1	4
DSJ_19170	ATP-dependent Clp endopeptidase, proteolytic subunit ClpP	CKS_0219	*clpP*	1	0
DSJ_19195	cytochrome o ubiquinol oxidase subunit II	CKS_0224	*cyoA*	0	3
DSJ_19205	cytochrome o ubiquinol oxidase subunit III	CKS_0226	*cyoC*	0	3
DSJ_19210	**cytochrome o ubiquinol oxidase subunit IV**	**CKS_0227**	***cyoD***	**0**	**0**
DSJ_19215	protoheme IX farnesyltransferase	CKS_0228	*cyoE*	0	2
DSJ_19245	**exodeoxyribonuclease VII small subunit**	**CKS_0234**	***xseB***	**0**	**0**
DSJ_19250	(2E,6E)-farnesyl diphosphate synthase	CKS_0235	*ispA*	3	2
DSJ_19265	**hypothetical protein**			**0**	**0**
DSJ_19295	nusB	CKS_0244	*nusB*	2	1
DSJ_19300	6,7-dimethyl-8-ribityllumazine synthase	CKS_0245	*ribE*	0	2
DSJ_19305	riboflavin biosynthesis protein RibD	CKS_0246	*ribD*	1	1
DSJ_19325	protein-export membrane protein SecD	CKS_0251	*secD*	0	6
DSJ_19330	preprotein translocase subunit YajC	CKS_0252	*yajC*	0	1
DSJ_19385	ribonuclease HI	CKS_0263	*rnhA*	0	1
DSJ_19390	DNA polymerase III subunit epsilon	CKS_0264	*dnaQ*	1	4
DSJ_19545	**group II intron reverse transcriptase/maturase**			**0**	**0**
DSJ_19550	**hypothetical protein**			**0**	**0**
DSJ_19600	phosphoheptose isomerase	CKS_0307	*lpcA*	1	0
DSJ_19785	**hypothetical protein**			**0**	**0**
DSJ_19895	hypothetical protein			5	3
DSJ_19930	hypothetical protein	CKS_0382		1	5
DSJ_20185	hypothetical protein			2	2.5
DSJ_20320	TonB system transport protein ExbD	CKS_0466	*exbD*	1	3
DSJ_20325	tonB-system energizer ExbB	CKS_0467	*exbB*	0	1
DSJ_20425	1-acyl-sn-glycerol-3-phosphate acyltransferase	CKS_0490	*plsC*	3	3
DSJ_20430	DNA topoisomerase IV subunit A	CKS_0491	*parC*	5	4
DSJ_20485	**hypothetical protein**			**0**	**0**
DSJ_20495	DNA topoisomerase IV subunit B	CKS_0507	*parE*	2	9

DSJ_20520	outer membrane channel protein TolC	CKS_0512	*tolC*	4	2
DSJ_20540	3,4-dihydroxy-2-butanone-4-phosphate synthase	CKS_0516	*ribB*	0	1
DSJ_20545	**hypothetical protein**	**CKS_0517**		**0**	**0**
DSJ_20570	**multifunctional CCA tRNA nucleotidyl transferase/2'3'-cyclic phosphodiesterase/2'nucleotidase/phosphatase**	**CKS_0522**	***cca***	**0**	**0**
DSJ_20580	dihydroneopterin aldolase	CKS_0524	*folB*	4	8
DSJ_20590	tRNA (adenosine(37)-N6)-threonylcarbamoyltransferase complex transferase subunit TsaD	CKS_0526	*ygjD*	0	1
DSJ_20595	30S ribosomal protein S21	CKS_0527	*rpsU*	0	1
DSJ_20600	DNA primase	CKS_0528	*dnaG*	0	4
DSJ_20605	**RNA polymerase sigma factor RpoD**	**CKS_0530**	***rpoD***	**0**	**0**
DSJ_20675	hypothetical protein	CKS_0544		2	4
DSJ_20680	hypothetical protein	CKS_0545		1	0
DSJ_20815	GGDEF domain-containing protein			4	6
DSJ_20870	**hypothetical protein**	**CKS_0586**		**0**	**0**
DSJ_20890	hypothetical protein	CKS_0591		0	1
DSJ_21010	**hypothetical protein**	**CKS_0623**		**0**	**0**
DSJ_21100	hypothetical protein			5	5
DSJ_21425	hypothetical protein			5	4
DSJ_21460	inorganic pyrophosphatase	CKS_0723	*ppa*	0	1
DSJ_21535	30S ribosomal protein S18	CKS_0742	*rpsR*	0	2
DSJ_21540	**primosomal replication protein N**	**CKS_0743**		**0**	**0**
DSJ_21545	30S ribosomal protein S6	CKS_0744	*rpsF*	3	0
DSJ_21630	tRNA (adenosine(37)-N6)-threonylcarbamoyltransferase complex ATPase subunit type 1 TsaE	CKS_0761	*yjeE*	0	1
DSJ_21670	phosphatidylserine decarboxylase	CKS_0769	*psd*	3	2
DSJ_21750	rod shape-determining protein MreD	CKS_0786	*mreD*	1	3
DSJ_21760	rod shape-determining protein	CKS_0788	*mreB*	1	3
DSJ_21780	**type II 3-dehydroquinate dehydratase**	**CKS_0793**		**0**	**0**
DSJ_21785	acetyl-CoA carboxylase, biotin carboxyl carrier protein	CKS_0794	*accB*	1	1
DSJ_21790	acetyl-CoA carboxylase biotin carboxylase subunit	CKS_0795	*accC*	4	1
DSJ_21850	protoporphyrinogen oxidase	CKS_4489	*hemG*	0	1
DSJ_21885	3-octaprenyl-4-hydroxybenzoate decarboxylase	CKS_4482	*ubiD*	3	1
DSJ_21905	**twin arginine-targeting protein translocase TatB**	**CKS_4478**	***tatB***	**0**	**0**

DSJ_21910	Sec-independent protein translocase TatA	CKS_4477	*tatA*	0	2
DSJ_21915	ubiB	CKS_4476	*ubiB*	3	4
DSJ_21920	hypothetical protein	CKS_4475	*yigP*	1	0
DSJ_21925	bifunctional demethylmenaquinone methyltransferase/2-methoxy-6-polyprenyl-1,4-benzoquinol methylase	CKS_4474	*ubiE*	1	2
DSJ_22045	**hypothetical protein**			**0**	**0**
DSJ_22080	DUF484 family protein	CKS_4439	*yigA*	2	1
DSJ_22085	diaminopimelate epimerase	CKS_4438	*dapF*	6	6
DSJ_22090	hypothetical protein	CKS_4437		0	1
DSJ_22100	**hydroxymethylbilane synthase**	**CKS_4434**	***hemC***	**0**	**0**
DSJ_22205	transcription termination factor Rho	CKS_4414	*rho*	0	4
DSJ_22450	glycine--tRNA ligase subunit beta	CKS_4359	*glyS*	1	0
DSJ_22455	glycine--tRNA ligase subunit alpha	CKS_4358	*glyQ*	3	1
DSJ_22480	tRNA uridine-5-carboxymethylaminomethyl(34) synthesis GTPase MnmE	CKS_4352	*mnmE*	3	4
DSJ_22485	membrane protein insertase YidC	CKS_4351	*yidC*	2	4
DSJ_22490	ribonuclease P protein component	CKS_4350	*rnpA*	2	0

[a] 168 genes with 0 read count (RC) are bold; [b] Locus tag and annotation names are from the complete *P. stewartii* DC283 genome (NCBI accession: CP017581); [c] CKS_locus tag and gene names re from the incomplete *P. stewartii* DC283 genome (NCBI accession: NZ_ AHIE00000000.1); [d] number of RC in LB medium

Table C.2. Genes from pDSJ010 with less than ten read counts in library grown in LB medium

Locus_tag[a]	Annotation name[a]	LB-1[b]	LB-2[b]
DSJ_25695	hypothetical protein	0	1
DSJ_25410	4-amino-4-deoxy-L-arabinose-phospho-UDP flippase	1	0
DSJ_25640	DNA-binding protein	1	2
DSJ_26385	hypothetical protein	1	6
DSJ_25525	hypothetical protein	2	7
DSJ_25340	hemolysin	3	3
DSJ_25810	hypothetical protein	3	0
DSJ_25415	4-amino-4-deoxy-L-arabinose lipid A transferase	4	3
DSJ_25835	chromosome partitioning protein ParB	6	6

[a] Locus tag and annotation names are from the complete *P. stewartii* DC283 genome (pDSJ010, NCBI accession: CP017592); [b] number of RC in LB medium

Table C.3. Genes with mutations two-fold differentially present in M9 minimal medium compared to LB medium

Locus_tag[a]	Annotation name[a]	CKS_#[b]	Gene[b]	Ratio1[c]	Ratio2[c]
DSJ_00220	type I glutamate--ammonia ligase	CKS_4298	*glnA*	0.11	0.04
DSJ_02575	glutaredoxin 3	CKS_1112	*grxC*	0.11	0.03
DSJ_04745	D-3-phosphoglycerate dehydrogenase	CKS_1393	*serA*	0.12	0.10
DSJ_01800	dihydroxy-acid dehydratase	CKS_1042	*ilvD*	0.13	0.08
DSJ_18335	cysteine synthase A	CKS_0047	*cysK*	0.15	0.50
DSJ_02585	glycerol-3-phosphate dehydrogenase	CKS_1110	*gpsA*	0.15	0.19
DSJ_01805	branched chain amino acid aminotransferase	CKS_1043	*ilvE*	0.15	0.09
DSJ_02775	argininosuccinate synthase	CKS_1063		0.15	0.10
DSJ_03880	glutamate synthase large subunit	CKS_4867		0.16	0.08
DSJ_01870	bifunctional phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase	CKS_0001	*purH*	0.16	0.08
DSJ_02755	phosphoenolpyruvate carboxylase	CKS_1067	*ppc*	0.16	0.07
DSJ_05180	phosphoadenosine phosphosulfate reductase	CKS_1302	*cysH*	0.17	0.23
DSJ_05965	homoserine kinase	CKS_5108	*thrB*	0.17	0.14
DSJ_02760	acetylornithine deacetylase	CKS_1066	*argE*	0.19	0.16
DSJ_02780	argininosuccinate lyase	CKS_1062	*argH*	0.19	0.09
DSJ_01785	ketol-acid reductoisomerase	CKS_1039	*ilvC*	0.20	0.11
DSJ_02770	acetylglutamate kinase	CKS_1064	*argB*	0.21	0.15
DSJ_06045	carbamoyl phosphate synthase small subunit	CKS_5128	*carA*	0.21	0.08
DSJ_06155	3-isopropylmalate dehydrogenase	CKS_5151	*leuB*	0.21	0.11
DSJ_04465	phosphoserine phosphatase SerB	CKS_4994	*serB*	0.21	0.06
DSJ_07685	phosphoribosylaminoimidazolesuccinocarboxamide synthase	CKS_1732	*purC*	0.22	0.14
DSJ_21585	adenylosuccinate synthase	CKS_0752	*purA*	0.22	0.06
DSJ_03875	sulfite reductase subunit alpha	CKS_4866		0.23	0.10
DSJ_04230	ornithine carbamoyltransferase	CKS_4942	*argI*	0.24	0.17
DSJ_02765	N-acetyl-gamma-glutamyl-phosphate reductase	CKS_1065	*argC*	0.24	0.27
DSJ_19690	glutamate-5-semialdehyde dehydrogenase	CKS_0327	*proA*	0.24	0.19
DSJ_06150	3-isopropylmalate dehydratase large subunit	CKS_5150	*leuC*	0.24	0.09
DSJ_15890	imidazole glycerol phosphate synthase subunit HisF	CKS_2290	*hisF*	0.24	0.17
DSJ_06160	2-isopropylmalate synthase	CKS_5152	*leuA*	0.24	0.12

DSJ_01810	acetolactate synthase 2 small subunit	CKS_1044		0.24	0.14
DSJ_01875	phosphoribosylamine--glycine ligase	CKS_0002	*purD*	0.25	0.10
DSJ_05960	bifunctional aspartate kinase/homoserine dehydrogenase I	CKS_5107	*thrA*	0.26	0.13
DSJ_15875	bifunctional imidazole glycerol-phosphate dehydratase/histidinol phosphatase	CKS_2293	*hisB*	0.26	0.15
DSJ_07560	ribosome biogenesis GTPase Der	CKS_1535	*der*	0.26	0.25
DSJ_07400	phosphoribosylformylglycinamidine synthase	CKS_1502	*purL*	0.26	0.11
DSJ_15870	histidinol-phosphate transaminase	CKS_2294	*hisC*	0.28	0.14
DSJ_06050	carbamoyl phosphate synthase large subunit	CKS_5129	*carB*	0.28	0.13
DSJ_05970	threonine synthase	CKS_5109	*thrC*	0.28	0.17
DSJ_05465	chorismate mutase	CKS_1239	*pheA*	0.28	0.24
DSJ_01225	GNAT family N-acetyltransferase	CKS_0924		0.28	0.32
DSJ_04895	diaminopimelate decarboxylase	CKS_1361	*lysA*	0.28	0.23
DSJ_05950	two-component system response regulator ArcA	CKS_5103	*arcA*	0.30	0.47
DSJ_07800	phosphoribosylformylglycinamidine cyclo-ligase	CKS_1706	*purM*	0.30	0.15
DSJ_19685	glutamate 5-kinase	CKS_0326	*proB*	0.31	0.16
DSJ_05175	sulfite reductase subunit beta	CKS_1303	*cysI*	0.31	0.19
DSJ_15880	imidazole glycerol phosphate synthase, glutamineamidotransferase subunit	CKS_2292	*hisH*	0.32	0.15
DSJ_17210	amidophosphoribosyltransferase	CKS_2047	*purF*	0.32	0.13
DSJ_15885	1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino]imidazole-4- carboxamide isomerase	CKS_2291	*hisA*	0.33	0.20
DSJ_04980	amino-acid N-acetyltransferase	CKS_1342	*argA*	0.33	0.31
DSJ_14615	septation protein A	CKS_2763	*yciB*	0.33	0.22
DSJ_10500	dihydroorotase	CKS_4192	*pyrC*	0.34	0.13
DSJ_01795	PLP-dependent threonine dehydratase	CKS_1041	*ilvA*	0.35	0.22
DSJ_15860	ATP phosphoribosyltransferase	CKS_2296	*hisG*	0.35	0.17
DSJ_08150	asnB	CKS_3751	*asnB*	0.36	0.28
DSJ_18705	5-(carboxyamino)imidazole ribonucleotide synthase	CKS_0120	*purK*	0.37	0.32
DSJ_14565	anthranilate synthase component I	CKS_2773	*trpE*	0.38	0.23
DSJ_02590	serine O-acetyltransferase	CKS_1109	*cysE*	0.38	0.20
DSJ_02410	orotate phosphoribosyltransferase	CKS_1152	*pyrE*	0.38	0.24
DSJ_00820	TriB protein	CKS_0815		0.39	0.15

DSJ_05170	sulfite reductase [NADPH] flavoprotein, alpha-component	CKS_1304	*cysJ*	0.41	0.22
DSJ_06145	3-isopropylmalate dehydratase small subunit	CKS_5149	*leuD*	0.41	0.21
DSJ_14580	bifunctional indole-3-glycerol phosphate synthase/phosphoribosylanthranilate isomerase	CKS_2770		0.41	0.22
DSJ_14570	anthranilate synthase component II	CKS_2772		0.41	0.19
DSJ_06335	nicotinate-nucleotide diphosphorylase	CKS_3613		0.42	0.47
DSJ_05265	adenylyl-sulfate kinase	CKS_1281	*cysC*	0.42	0.31
DSJ_14585	tryptophan synthase subunit beta	CKS_2769	*trpB*	0.42	0.07
DSJ_14465	orotidine 5'-phosphate decarboxylase	CKS_2797	*pyrF*	0.42	0.12
DSJ_00865	bifunctional succinylornithine transaminase/acetylornithine transaminase	CKS_0821	*argD*	0.42	0.50
DSJ_15895	bifunctional phosphoribosyl-AMP cyclohydrolase/phosphoribosyl-ATP diphosphatase	CKS_2289	*hisI*	0.43	0.17
DSJ_15865	histidinol dehydrogenase	CKS_2295	*hisD*	0.43	0.18
DSJ_14590	tryptophan synthase subunit alpha	CKS_2768	*trpA*	0.44	0.23
DSJ_02940	homoserine O-succinyltransferase	CKS_4655	*metA*	0.45	0.40
DSJ_08500	hypothetical protein	CKS_3842		0.47	0.29
DSJ_04740	ribose 5-phosphate isomerase A	CKS_1394	*rpiA*	0.47	0.10
DSJ_14575	anthranilate phosphoribosyltransferase	CKS_2771		0.47	0.15
DSJ_09690	dihydroorotate dehydrogenase (quinone)	CKS_4024	*pyrD*	0.48	0.20
DSJ_06890	transcriptional repressor MprA	CKS_1456	*mprA*	0.50	0.44
DSJ_02055[d]	transcriptional regulator RbsR	CKS_1231	*rbsR*	2.12	2.22
DSJ_25325[e]	hypothetical protein			2.23	2.07

[a] Locus tag and annotation names are from the complete *P. stewartii* DC283 genome (NCBI accession: CP017581); [b] CKS_locus tag and gene names are from the incomplete *P. stewartii* DC283 genome (NCBI accession: NZ_ AHIE00000000.1); [c] ratio of number of RC in M9 minimal medium over number of RC in LB medium; [d] gene with better fitness in M9 medium when being inserted; [e] gene from pDSJ010 (NCBI accession: CP017592).

Table C.4. Genes with mutations ten-fold differentially present *in planta* growth compared to LB medium with greater than 100 read counts in each LB sample[a]

Locus_tag[b]	Annotation name[b]	CKS_#[c]	Gene[c]	Ratio1[d]	Ratio2[d]
DSJ_17215	**colicin V production protein**	**CKS_2046**	***cvpA***	**0.00**	**0.00**
DSJ_09680	**glycosyl transferase**	**CKS_4022**	***wceO***	**0.01**	**0.00**
DSJ_17620	**hypothetical protein**	**CKS_1961**		**0.03**	**0.00**
DSJ_16075	polysaccharide export protein Wza	CKS_2244	*wza*	0.09	0.00
DSJ_09635	**porin OmpC**	**CKS_4012**	***ompF***	**0.00**	**0.00**
DSJ_02405	**ribonuclease PH**	**CKS_1153**		**0.01**	**0.01**
DSJ_10540	**anti-sigma-28 factor FlgM**	**CKS_4201**	***flgM***	**0.01**	**0.01**
DSJ_09745	**porin OmpA**	**CKS_4036**	***ompA***	**0.01**	**0.00**
DSJ_03430	**SgbH**	**CKS_4767**	***yiaQ***	**0.03**	**0.01**
DSJ_00740	**hypothetical protein**	**CKS_0834**		**0.02**	**0.00**
DSJ_00315	**hypothetical protein**	**CKS_1487**		**0.03**	**0.01**
DSJ_01080	**rhomboid family intramembrane serine protease GlpG**	**CKS_0889**	***glpG***	**0.04**	**0.01**
DSJ_10435	**MdoG**	**CKS_4179**		**0.00**	**0.00**
DSJ_03910	**ribosomal subunit interface protein**	**CKS_4873**		**0.01**	**0.01**
DSJ_16055	**amylovoran biosynthesis protein AmsB**	**CKS_2248**	***wceB***	**0.01**	**0.01**
DSJ_00305	**shikimate dehydrogenase**	**CKS_1485**	***aroE***	**0.01**	**0.01**
DSJ_10440	**glucan biosynthesis glucosyltransferase H**	**CKS_4180**		**0.02**	**0.00**
DSJ_06950	peptide ABC transporter ATP-binding protein	CKS_1472		0.08	0.01
DSJ_07830	**exopolyphosphatase**	**CKS_1700**	***ppx***	**0.01**	**0.02**
DSJ_21615	**tRNA (adenosine(37)-N6)-dimethylallyltransferaseMiaA**	**CKS_0758**	***miaA***	**0.02**	**0.02**
DSJ_07805	**phosphoribosylglycinamide formyltransferase**	**CKS_1705**	***purN***	**0.01**	**0.01**
DSJ_04470	**DNA repair protein RadA**	**CKS_4995**	***radA***	**0.03**	**0.01**
DSJ_22095	**adenylate cyclase**	**CKS_4435**	***cyaA***	**0.06**	**0.02**
DSJ_18455	capsid protein	CKS_0075		0.08	0.02
DSJ_19160	**endopeptidase La**	**CKS_0217**	***lon***	**0.01**	**0.01**
DSJ_18350	**PTS glucose transporter subunit IIA**	**CKS_0050**	***crr***	**0.02**	**0.01**
DSJ_04620	**hypothetical protein**	**CKS_1418**	***yggT***	**0.07**	**0.02**
DSJ_03785	**FxsA**	**CKS_4847**	***fxsA***	**0.06**	**0.02**
DSJ_03050	**phosphate-starvation-inducible protein PsiE**	**CKS_4680**	***psiE***	**0.08**	**0.02**
DSJ_00170	**bifunctional GTP diphosphokinase/guanosine-3',5'-bis(diphosphate) 3'-diphosphatase**	**CKS_4309**	***spoT***	**0.07**	**0.03**
DSJ_17265	DNA-binding response regulator	CKS_2035		0.08	0.01

DSJ_05305	**peptidase**	**CKS_1273**	***nlpD***	**0.01**	**0.01**
DSJ_03865	**sspA**	**CKS_4864**	***sspA***	**0.03**	**0.02**
DSJ_08295	**citrate (Si)-synthase**	**CKS_3782**	***gltA***	**0.06**	**0.02**
DSJ_03780	**aspartate ammonia-lyase**	**CKS_4845**	***aspA***	**0.04**	**0.01**
DSJ_20150	hypothetical protein	CKS_0427		0.06	0.02
DSJ_05920	phosphoglycerate mutase	CKS_5097	*ytjC*	0.10	0.04
DSJ_02695	HslU--HslV peptidase ATPase subunit	CKS_1085	*hslU*	0.08	0.01
DSJ_10795	**sensor protein PhoQ**	**CKS_4261**	***phoQ***	**0.03**	**0.01**
DSJ_08530	**kinase inhibitor**	**CKS_3848**	***ybhB***	**0.03**	**0.02**
DSJ_13265	**hypothetical protein**	**CKS_3483**		**0.08**	**0.03**
DSJ_08810	peptidase	CKS_5689		0.10	0.06
DSJ_00255	protein disulfide oxidoreductase DsbA	CKS_4290	*dsbA*	0.07	0.03
DSJ_19055	ABC transporter ATP-binding protein	CKS_0194		0.08	0.03
DSJ_18990	recombination protein RecR	CKS_0180	*recR*	0.09	0.03
DSJ_04890	LysR family transcriptional regulator	CKS_1362	*lysR*	0.02	0.01
DSJ_14075	choline ABC transporter permease	CKS_2885		0.09	0.01
DSJ_04830	**hypothetical protein**	**CKS_1376**		**0.03**	**0.06**
DSJ_00995	**GMP/IMP nucleotidase**			**0.05**	**0.03**
DSJ_03435	xylulose 5-phosphate 3-epimerase	CKS_4768	*yiaR*	0.04	0.02
DSJ_04095	**lipoprotein NlpI**	**CKS_4911**	***nlpI***	**0.03**	**0.02**
DSJ_05575	**hypothetical protein**	**CKS_5024**		**0.07**	**0.04**
DSJ_04235	**ribonuclease E inhibitor B**	**CKS_4943**	***rraB***	**0.07**	**0.04**
DSJ_21580	transcriptional repressor NsrR	CKS_0751	*nsrR*	0.06	0.01
DSJ_22215	**ATP-dependent RNA helicase RhlB**	**CKS_4412**	***rhlB***	**0.02**	**0.02**
DSJ_15515	**phosphate starvation protein PhoH**	**CKS_1874**		**0.05**	**0.07**
DSJ_21690	**hypothetical protein**	**CKS_0773**	***yhcN***	**0.07**	**0.05**
DSJ_01165	**hypothetical protein**	**CKS_0912**		**0.10**	**0.07**
DSJ_21590	DUF2065 domain-containing protein	CKS_0753		0.08	0.04
DSJ_05720	shikimate kinase II	CKS_5053		0.10	0.05
DSJ_06400	UPF0231 family protein	CKS_5203	*yacL*	0.10	0.04
DSJ_08680	**L,D-transpeptidase**	**CKS_3918**		**0.01**	**0.02**
DSJ_10450	**secY/secA suppressor protein**	**CKS_4182**	***msyB***	**0.06**	**0.07**
DSJ_15525	**iron uptake system protein EfeO**	**CKS_1876**	***efeO***	**0.07**	**0.05**
DSJ_01890	hypothetical protein	CKS_0005	*yjaG*	0.06	0.04
DSJ_17435	transposase	CKS_1999		0.10	0.05
DSJ_21775	sulfoxide reductase heme-binding subunit YedZ	CKS_0792	*yedZ*	0.10	0.05
DSJ_04650	16S rRNA (uracil(1498)-N(3))-methyltransferase	CKS_1412		0.08	0.03
DSJ_01250	rsmD	CKS_0929		0.08	0.02
DSJ_19710	**hypothetical protein**	**CKS_0331**		**0.05**	**0.05**
DSJ_18650	hypothetical protein	CKS_0110		0.09	0.04

DSJ_04750	5-formyltetrahydrofolate cyclo-ligase	CKS_1392	*ygfA*	0.04	0.03
DSJ_07700	glycine cleavage system transcriptional repressor	CKS_1729	*gcvR*	0.04	0.03
DSJ_05455	3-deoxy-7-phosphoheptulonate synthase	CKS_1241	*aroF*	0.09	0.02
DSJ_02470	putative lipopolysaccharide heptosyltransferase III	CKS_1139		0.05	0.06
DSJ_02415	nucleoid occlusion factor SlmA	CKS_1151		0.03	0.03
DSJ_05480	23S rRNA pseudouridine(1911/1915/1917) synthase	CKS_1235	*rluD*	0.03	0.04
DSJ_17390	thiol:disulfide interchange protein	CKS_2008	*ccmG*	0.07	0.04
DSJ_00765	hypothetical protein			0.09	0.07
DSJ_08570	hypothetical protein	CKS_3857	*ybhK*	0.07	0.05
DSJ_14815	hypothetical protein	CKS_2718	*yeaE*	0.10	0.07
DSJ_21285	penicillin-binding protein activator	CKS_0682	*yraM*	0.05	0.03
DSJ_19990	hypothetical protein	CKS_0393		0.08	0.03
DSJ_08330	succinate--CoA ligase subunit beta	CKS_3791	*sucC*	0.10	0.05
DSJ_02685	septal ring assembly protein ZapB	CKS_1087		0.05	0.06
DSJ_22220	guanosine-5'-triphosphate,3'-diphosphate pyrophosphatase	CKS_4411	*gpp*	0.03	0.02
DSJ_17105	hypothetical protein	CKS_2070	*yfbV*	0.07	0.06
DSJ_13735	hypothetical protein			0.09	0.07
DSJ_16065	tyrosine-protein kinase	CKS_2246	*wzc*	0.04	0.03
DSJ_17710	hypothetical protein			0.07	0.06
DSJ_04500	MrkB			0.08	0.06
DSJ_12995	murein lipoprotein	CKS_3420	*lpp*	0.09	0.06
DSJ_01150	quercetin 2,3-dioxygenase	CKS_0908	*yhhW*	0.09	0.05
DSJ_05445	hypothetical protein	CKS_1243		0.08	0.05
DSJ_09790	heat-shock protein HspQ	CKS_4046	*hspQ*	0.09	0.04
DSJ_07550	hypothetical protein	CKS_1533	*yfgM*	0.08	0.02
DSJ_19145	hypothetical protein	CKS_0214		0.10	0.04
DSJ_01085	thiosulfate sulfurtransferase	CKS_0890	*glpE*	0.07	0.02
DSJ_21320	2-dehydro-3-deoxyphosphooctonate aldolase			0.06	0.06
DSJ_16815	hypothetical protein			0.10	0.09
DSJ_08960	osmoprotectant uptake system permease	CKS_3938	*yehY*	0.08	0.06
DSJ_16475	16S rRNA pseudouridine(516) synthase	CKS_2148	*rsuA*	0.09	0.07
DSJ_00110	DNA-3-methyladenine glycosidase	CKS_4324	*tag*	0.07	0.04
DSJ_03525	glycosyl hydrolase family 32	CKS_4790		0.04	0.02
DSJ_21890	transcriptional activator RfaH	CKS_4481	*rfaH*	0.04	0.10
DSJ_20970	amino acid-binding protein	CKS_0611		0.09	0.07

DSJ_08155	UMP phosphatase	CKS_3752	*nagD*	0.08	0.07
DSJ_11585	DNA-directed RNA polymerase subunit alpha	CKS_2498		0.10	0.08
DSJ_07450	3-phenylpropionic acid transporter	CKS_1511	*hcaT*	0.10	0.09
DSJ_07750	ABC transporter permease	CKS_1718		0.10	0.06
DSJ_12630	type VI secretion system, core protein	CKS_3339		0.09	0.05
DSJ_05660	branched-chain amino acid transport system II carrier protein	CKS_5039	*brnQ*	0.09	0.05
DSJ_18530	antitermination protein	CKS_0089		0.09	0.04
DSJ_04280	peptidase M20	CKS_4952		0.08	0.03
DSJ_06470	multicopper oxidase			0.08	0.02
DSJ_04460	hypothetical protein	CKS_4993	*ytjB*	0.04	0.02
DSJ_03410	3-dehydro-L-gulonate 2-dehydrogenase	CKS_4763	*yiaK*	0.05	0.03
DSJ_00920	phosphoglycolate phosphatase	CKS_0855	*gph*	0.08	0.05
DSJ_21180	transcriptional regulator ExuR	CKS_0661	*exuR*	0.10	0.05
DSJ_09140	23S rRNA (adenine(1618)-N(6))-methyltransferase			0.01	0.03
DSJ_20975	antibiotic biosynthesis monooxygenase	CKS_0612		0.07	0.07
DSJ_09975	antitermination protein	CKS_4080		0.10	0.09
DSJ_20230	GNAT family N-acetyltransferase	CKS_0444		0.09	0.10
DSJ_15800	GntP family transporter	CKS_2309	*ygbN*	0.09	0.05
DSJ_19510	amino acid ABC transporter permease	CKS_0289		0.10	0.04
DSJ_00125	LacI family transcriptional regulator	CKS_4321		0.07	0.10
DSJ_21745	septum formation inhibitor Maf	CKS_0785	*yhdE*	0.08	0.06
DSJ_02525	glycine C-acetyltransferase			0.09	0.09
DSJ_21575	ribonuclease R	CKS_0750	*rnr*	0.03	0.02
DSJ_06630	[protein-PII] uridylyltransferase	CKS_5256	*glnD*	0.06	0.02
DSJ_21595	protease modulator HflC	CKS_0754	*hflC*	0.07	0.09
DSJ_20440	hypothetical protein			0.09	0.08
DSJ_02080	D-ribose pyranase	CKS_1226	*rbsD*	0.10	0.09
DSJ_09465	thioredoxin-disulfide reductase	CKS_3976	*trxB*	0.09	0.04
DSJ_05040	23S rRNA (cytidine(2498)-2'-O)-methyltransferaseRlmM	CKS_1330	*ygdE*	0.10	0.06
DSJ_16035	colanic acid biosynthesis pyruvyl transferase WcaK	CKS_2252	*wceJ*	0.03	0.05
DSJ_01060	Fe-S biogenesis protein NfuA	CKS_0884		0.05	0.08
DSJ_03655	DNA-binding response regulator			0.09	0.08
DSJ_01665	potassium-transporting ATPase subunit KdpA	CKS_1013	*kdpA*	0.06	0.04
DSJ_09360	RumB			0.09	0.03
DSJ_20270	hypothetical protein	CKS_0453		0.08	0.10
DSJ_03215	ActP	CKS_4718	*actP*	0.09	0.07

DSJ_10160	hypothetical protein	CKS_4117		0.10	0.08
DSJ_00320	hypothetical protein	CKS_1488	*smg*	0.10	0.05
DSJ_15760	2-dehydropantoate 2-reductase	CKS_2319		0.09	0.05
DSJ_01015	two-component system sensor histidine kinase EnvZ	CKS_0875	*envZ*	0.05	0.02
DSJ_15940	UDP-glucose 4-epimerase GalE	CKS_2274		0.09	0.09
DSJ_09165	threonine/homoserine exporter RhtA			0.10	0.07
DSJ_06550	penicillin-binding protein 1B	CKS_5239	*mrcB*	0.05	0.02
DSJ_01235	cell division protein FtsX	CKS_0926	*ftsX*	0.10	0.02
DSJ_07485	Fe-S cluster assembly transcriptional regulator IscR	CKS_1518	*iscR*	0.05	0.09
DSJ_21740	ribonuclease E/G	CKS_0784	*rng*	0.08	0.06
DSJ_20045	hypothetical protein	CKS_0404		0.06	0.09
DSJ_18955	LysR family transcriptional regulator	CKS_0171		0.09	0.09
DSJ_07555	outer membrane protein assembly factor BamB	CKS_1534		0.10	0.04
DSJ_21510	cell envelope opacity-associated protein A YtfB	CKS_0734		0.06	0.07
DSJ_08550	malonyl-[acyl-carrier protein] O-methyltransferase BioC	CKS_3852	*bioC*	0.09	0.09
DSJ_00705	hypothetical protein			0.10	0.06
DSJ_01590	spermidine/putrescine ABC transporter permease	CKS_0997		0.10	0.04
DSJ_00360	16S rRNA (cytosine(967)-C(5))-methyltransferase	CKS_1556	*rsmB*	0.06	0.08
DSJ_00590	peptidylprolyl isomerase	CKS_1603	*slyD*	0.06	0.02
DSJ_02195	hypothetical protein			0.08	0.05
DSJ_12635	type VI secretion-associated protein	CKS_3340		0.07	0.05
DSJ_01755	short-chain dehydrogenase	CKS_1033		0.09	0.07
DSJ_07905	LacI family transcriptional regulator	CKS_3702		0.09	0.06
DSJ_21565	isovaleryl-CoA dehydrogenase	CKS_0748	*aidB*	0.05	0.03
DSJ_01035	RNA-binding transcriptional accessory protein	CKS_0879	*yhgF*	0.10	0.04
DSJ_08850	hypothetical protein	CKS_5680		0.10	0.05
DSJ_02985	transcriptional regulator IclR	CKS_4668	*iclR*	0.08	0.06
DSJ_01455	16S rRNA (guanine(1516)-N(2))-methyltransferase	CKS_0972	*yhiQ*	0.09	0.09
DSJ_20155	hypothetical protein	CKS_0428		0.10	0.06
DSJ_03450	Exoz	CKS_4771		0.09	0.07
DSJ_17910	transcriptional regulator			0.09	0.08
DSJ_03645	LysR family transcriptional regulator	CKS_4817		0.10	0.05
DSJ_16410	1-phosphofructokinase	CKS_2164	*fruK*	0.08	0.06
DSJ_09400	N-acetylmuramoyl-L-alanine amidase			0.08	0.07

DSJ_22255	alpha/beta hydrolase	CKS_4404		0.09	0.06
DSJ_19850	LysR family transcriptional regulator	CKS_0364		0.08	0.06
DSJ_04850	single-stranded-DNA-specific exonuclease RecJ	CKS_1372	*recJ*	0.09	0.05
DSJ_06190	DNA-binding transcriptional regulator FruR	CKS_5158	*fruR*	0.07	0.07
DSJ_21380	aldose epimerase	CKS_0706		0.07	0.05
DSJ_16260	GNAT family N-acetyltransferase			0.09	0.08
DSJ_22110	uroporphyrinogen-III C-methyltransferase	CKS_4432	*hemX*	0.05	0.04
DSJ_20700	5'-nucleotidase, lipoprotein e(P4) family	CKS_0550		0.07	0.05
DSJ_02110	hypothetical protein	CKS_1219	*viaA*	0.07	0.06
DSJ_15465	two-component system sensor histidine kinase BasS	CKS_1863	*basS*	0.10	0.09
DSJ_02620	divalent metal cation transporter FieF	CKS_1102	*fieF*	0.06	0.04
DSJ_21435	fructose-bisphosphatase	CKS_0718	*fbp*	0.06	0.05
DSJ_01300	inositol 2-dehydrogenase	CKS_0940		0.07	0.06
DSJ_01815	acetolactate synthase 2 catalytic subunit	CKS_1045		0.09	0.10
DSJ_02895	excinuclease ABC subunit A	CKS_4643	*uvrA*	0.10	0.05
DSJ_01115	1,4-alpha-glucan branching enzyme			0.07	0.04
DSJ_02050	MFS transporter	CKS_1232	*hsrA*	0.10	0.08

[a] 47 genes with less than 10 read counts (RC) in plant samples (Table 3) were bold; [b] Locus tag and annotation names were used from the complete *P. stewartii* DC283 genome (NCBI accession: CP017581); [c] CKS_locus tag and gene names were used from the incomplete *P. stewartii* DC283 genome (NCBI accession: NZ_ AHIE00000000.1); [d] ratio of number of RC *in planta* growth over number of RC in LB medium

Table C.5. Genes with mutations ten-fold differentially present *in planta* growth compared to LB medium with read counts between 10-100 in each LB sample[a]

Locus_tag[b]	Annotation name[b]	CKS_#[c]	Gene[c]	Ratio1[d]	Ratio2[d]
DSJ_02860	**IS630 family transposase**			**0.00**	**0.00**
DSJ_03320	**iron transporter**	**CKS_4743**		**0.00**	**0.00**
DSJ_03905	**PTS IIA-like nitrogen regulatory protein PtsN**	**CKS_4872**	***ptsN***	**0.00**	**0.00**
DSJ_04770	**Xaa-Pro aminopeptidase**	**CKS_1389**	***pepP***	**0.00**	**0.00**
DSJ_06075	**ksgA**	**CKS_5134**		**0.00**	**0.00**
DSJ_06085	**peptidylprolyl isomerase SurA**	**CKS_5136**	***surA***	**0.00**	**0.00**
DSJ_06365	**transcriptional regulator PdhR**	**CKS_5195**	***pdhR***	**0.00**	**0.00**
DSJ_06395	**bifunctional aconitate hydratase 2/2-methylisocitrate dehydratase**	**CKS_5202**	***acnB***	**0.00**	**0.00**
DSJ_06860	**thioredoxin TrxC**			**0.00**	**0.00**
DSJ_08165	**N-acetylglucosamine-6-phosphate deacetylase**	**CKS_3754**	***nagA***	**0.00**	**0.00**
DSJ_08310	**succinate dehydrogenase flavoprotein subunit**	**CKS_3786**	***sdhA***	**0.00**	**0.00**
DSJ_08375	**cell envelope integrity protein TolA**	**CKS_3800**	***tolA***	**0.00**	**0.00**
DSJ_09685	**flippase**	**CKS_4023**	***wzx2***	**0.00**	**0.00**
DSJ_10155	**hypothetical protein**	**CKS_4116**		**0.00**	**0.00**
DSJ_10665	**fabG**	**CKS_4230**	***fabG***	**0.00**	**0.00**
DSJ_16045	**amylovoran biosynthesis protein AmsD**	**CKS_2250**	***wceN***	**0.00**	**0.00**
DSJ_16070	**protein tyrosine phosphatase**	**CKS_2245**	***wzb***	**0.00**	**0.00**
DSJ_16855	**hypothetical protein**			**0.00**	**0.00**
DSJ_17020	**NADH-quinone oxidoreductase subunit M**	**CKS_2087**	***nuoM***	**0.00**	**0.00**
DSJ_17060	**NADH-quinone oxidoreductase subunit E**	**CKS_2079**	***nuoE***	**0.00**	**0.00**
DSJ_17075	**NADH-quinone oxidoreductase subunit A**	**CKS_2076**	***nuoA***	**0.00**	**0.00**
DSJ_18780	**hypothetical protein**	**CKS_0137**		**0.00**	**0.00**
DSJ_18870	**phage tail protein**	**CKS_0155**		**0.00**	**0.00**
DSJ_21300	**osmotically-inducible protein OsmY**	**CKS_0685**	***yraP***	**0.00**	**0.00**
DSJ_21680	**malate dehydrogenase**	**CKS_0771**	***mdh***	**0.00**	**0.00**
DSJ_22350	**MarR family transcriptional regulator**	**CKS_4382**		**0.00**	**0.00**
DSJ_02350	hypothetical protein	CKS_1163		0.00	0.05
DSJ_03900	RNase adaptor protein RapZ	CKS_4871	*yhbJ*	0.00	0.03
DSJ_09040	hypothetical protein	CKS_3956		0.00	0.06
DSJ_09675	NAD(P)H-dependent FMN reductase	CKS_4021	*ssuE*	0.00	0.06
DSJ_14785	L-asparaginase 1	CKS_2724	*ansA*	0.00	0.03

DSJ_15935	phosphogluconate dehydrogenase (NADP(+)-dependent, decarboxylating)	CKS_2276	*gnd*	0.00	0.05
DSJ_17025	NADH-quinone oxidoreductase subunit L	CKS_2086	*nuoL*	0.00	0.03
DSJ_19290	thiamine-phosphate kinase	CKS_0243	*thiL*	0.00	0.04
DSJ_21420	hypothetical protein	CKS_0714	*yjgA*	0.00	0.07
DSJ_22170	O-antigen translocase	CKS_4423	*wzxE*	0.00	0.03
DSJ_03455	ligand-gated channel protein	CKS_4772		0.00	0.04
DSJ_06325	prepilin peptidase-dependent pilin	CKS_5186	*ppdD*	0.00	0.09
DSJ_15905	ABC transporter substrate-binding protein			0.00	0.05
DSJ_16060	amylovoran biosynthesis protein AmsC	CKS_2247	*wceL*	0.01	0.00
DSJ_06070	Co2+/Mg2+ efflux protein ApaG	CKS_5133	*apaG*	0.01	0.01
DSJ_04910	hypothetical protein	CKS_1358	*ygdQ*	0.01	0.01
DSJ_17080	transcriptional regulator LrhA	CKS_2075	*lrhA*	0.01	0.03
DSJ_16030	colanic acid biosynthesis glycosyltransferase WcaL	CKS_2253	*wceK*	0.02	0.03
DSJ_22065	DNA helicase II	CKS_4442	*uvrD*	0.02	0.09
DSJ_21060	MrkB			0.02	0.02
DSJ_21465	gamma-glutamylcyclotransferase	CKS_0724	*ytfP*	0.02	0.04
DSJ_08385	peptidoglycan-associated lipoprotein	CKS_3802	*pal*	0.02	0.03
DSJ_18345	phosphoenolpyruvate--protein phosphotransferase	CKS_0049	*ptsI*	0.02	0.03
DSJ_04100	DEAD/DEAH family ATP-dependent RNA helicase	CKS_4912	*deaD*	0.02	0.07
DSJ_20085	hypothetical protein	CKS_0412		0.02	0.09
DSJ_16355	iron ABC transporter ATP-binding protein	CKS_2175		0.03	0.01
DSJ_01075	DeoR/GlpR family transcriptional regulator	CKS_0888	*glpR*	0.03	0.01
DSJ_07080	ATP-dependent RNA helicase SrmB	CKS_1669	*srmB*	0.03	0.03
DSJ_06080	4-hydroxythreonine-4-phosphate dehydrogenase PdxA	CKS_5135	*pdxA*	0.03	0.00
DSJ_16570	phage tail protein			0.03	0.09
DSJ_04020	RNA-binding protein	CKS_4895		0.03	0.03
DSJ_15675	sulfate transporter			0.03	0.09
DSJ_19070	Hha toxicity attenuator	CKS_0197		0.03	0.00
DSJ_14735	patatin family protein	CKS_2735	*rssA*	0.03	0.01
DSJ_07245	hypothetical protein	CKS_1691		0.03	0.09
DSJ_20805	purine-nucleoside phosphorylase	CKS_0572	*deoD*	0.03	0.08
DSJ_10010	hypothetical protein	CKS_4089		0.03	0.04
DSJ_02295	hypothetical protein			0.03	0.00
DSJ_00965	hypothetical protein			0.03	0.07

DSJ_04140	protease	CKS_4920	*yhbO*	0.03	0.10
DSJ_04875	cupin			0.03	0.03
DSJ_14355	phage shock protein PspA	CKS_2818		0.03	0.07
DSJ_17050	NADH-quinone oxidoreductase subunit G	CKS_2081	*nuoG*	0.04	0.04
DSJ_17065	NADH-quinone oxidoreductase subunit C/D	CKS_2078	*nuoC*	0.04	0.00
DSJ_09025	hypothetical protein	CKS_3953		0.04	0.10
DSJ_15080	C-terminal processing peptidase	CKS_1779	*prc*	0.04	0.00
DSJ_21685	arginine repressor	CKS_0772	*argR*	0.04	0.06
DSJ_02545	lipopolysaccharide heptosyltransferase 1	CKS_1118	*rfaC*	0.04	0.00
DSJ_00215	GTP-binding protein TypA	CKS_4300	*typA*	0.04	0.00
DSJ_17070	NADH dehydrogenase	CKS_2077	*nuoB*	0.04	0.01
DSJ_14910	septum site-determining protein MinD	CKS_2696	*minD*	0.04	0.06
DSJ_10180	hypothetical protein	CKS_4121		0.04	0.00
DSJ_04645	glutathione synthase	CKS_1413	*gshB*	0.04	0.00
DSJ_06375	aceF	CKS_5197	*aceF*	0.04	0.04
DSJ_07135	DNA repair protein RecO	CKS_1657	*recO*	0.04	0.04
DSJ_04625	YggS family pyridoxal phosphate enzyme	CKS_1417	*yggS*	0.04	0.07
DSJ_19075	transcriptional regulator	CKS_0198		0.04	0.04
DSJ_20080	phage baseplate protein	CKS_0411		0.04	0.06
DSJ_09180	transcriptional regulator MntR			0.04	0.04
DSJ_19530	OHCU decarboxylase	CKS_0293		0.04	0.07
DSJ_16290	cytidine deaminase	CKS_2189	*cdd*	0.04	0.07
DSJ_08795	hypothetical protein	CKS_5692		0.04	0.01
DSJ_19155	DNA-binding protein HU	CKS_0216	*hupB*	0.04	0.03
DSJ_09800	acylphosphatase	CKS_4048	*yccX*	0.05	0.09
DSJ_02280	peroxiredoxin			0.05	0.05
DSJ_04410	cys-tRNA(pro)/cys-tRNA(cys) deacylase	CKS_4982		0.05	0.09
DSJ_08030	nicotinic acid mononucleotide adenylyltransferase	CKS_3732	*nadD*	0.05	0.00
DSJ_00310	L-threonylcarbamoyladenylate synthase type 1 TsaC	CKS_1486		0.05	0.05
DSJ_03460	lysine 6-monooxygenase	CKS_4773		0.05	0.02
DSJ_17700	hypothetical protein	CKS_1945		0.05	0.10
DSJ_09420	hypothetical protein			0.05	0.00
DSJ_22425	glutathione S-transferase	CKS_4365	*yibF*	0.05	0.05
DSJ_02605	DNA-binding response regulator	CKS_1106	*cpxR*	0.05	0.03
DSJ_14995	hypothetical protein	CKS_2678	*yobD*	0.05	0.03
DSJ_04215	reactive intermediate/imine deaminase	CKS_4938	*ridA*	0.05	0.03

DSJ_20930	hypothetical protein	CKS_0600		0.05	0.04
DSJ_07530	type IV pilus biogenesis/stability protein PilW	CKS_1529		0.05	0.05
DSJ_18845	phage tail protein	CKS_0150		0.05	0.00
DSJ_13295	electron transport complex subunit RsxB	CKS_3489	*rsxB*	0.05	0.08
DSJ_21345	cytoplasmic protein	CKS_0699		0.05	0.08
DSJ_09315	sensory transduction regulator			0.05	0.04
DSJ_18850	phage tail protein	CKS_0151		0.06	0.00
DSJ_04720	arginine transporter	CKS_1398	*argO*	0.06	0.06
DSJ_00595	hypothetical protein	CKS_1604	*yheV*	0.06	0.03
DSJ_22060	magnesium and cobalt transport protein CorA	CKS_4443	*corA*	0.06	0.08
DSJ_04090	polyribonucleotide nucleotidyltransferase	CKS_4910	*pnp*	0.06	0.00
DSJ_04615	non-canonical purine NTP pyrophosphatase, RdgB/HAM1 family	CKS_1419		0.06	0.02
DSJ_18135	transcriptional regulator	CKS_3632		0.06	0.02
DSJ_04825	hypothetical protein	CKS_1377	*ygfY*	0.06	0.09
DSJ_04950	hypothetical protein	CKS_1349		0.06	0.06
DSJ_13040	LysR family transcriptional regulator	CKS_3433	*ydhB*	0.06	0.00
DSJ_20455	PTS fructose transporter subunit IIB	CKS_0498		0.06	0.00
DSJ_04920	RNA pyrophosphohydrolase	CKS_1356		0.06	0.04
DSJ_10105	hypothetical protein	CKS_4107		0.06	0.05
DSJ_05015	cysteine desulfurase, sulfur acceptor subunit CsdE	CKS_1335		0.06	0.00
DSJ_08765	hypothetical protein	CKS_5698		0.06	0.03
DSJ_06920	hypothetical protein	CKS_1465		0.06	0.06
DSJ_02845	DUF1471 domain-containing protein	CKS_5511		0.06	0.05
DSJ_20560	inorganic triphosphatase	CKS_0520		0.06	0.06
DSJ_05125	nucleoside triphosphate pyrophosphohydrolase	CKS_1312	*mazG*	0.06	0.02
DSJ_01140	gluconokinase	CKS_0906	*gntK*	0.06	0.04
DSJ_10715	histidine triad nucleotide-binding protein	CKS_4242	*hinT*	0.06	0.05
DSJ_02790	NAD(P)(+) transhydrogenase	CKS_1060	*sthA*	0.07	0.03
DSJ_22410	superoxide dismutase	CKS_4368	*sodA*	0.07	0.10
DSJ_12260	Arc family DNA-binding protein	CKS_2635		0.07	0.09
DSJ_08525	6-phosphogluconolactonase	CKS_3847	*pgl*	0.07	0.03
DSJ_07340	hypothetical protein	CKS_2955		0.07	0.06
DSJ_03415	hypothetical protein	CKS_4764		0.07	0.00
DSJ_13105	hypothetical protein	CKS_3447		0.07	0.07
DSJ_02735	bifunctional aspartate	CKS_1077	*metL*	0.07	0.06

	kinase/homoserine dehydrogenase II				
DSJ_04295	ABC transporter permease	CKS_4955		0.07	0.07
DSJ_01625	hypothetical protein	CKS_1005		0.07	0.06
DSJ_00175	tRNA (guanosine(18)-2'-O)-methyltransferase TrmH	CKS_4308	*trmH*	0.07	0.06
DSJ_14050	hypothetical protein	CKS_2890	*ynfB*	0.07	0.02
DSJ_13590	modification methylase			0.07	0.04
DSJ_17055	NADH-quinone oxidoreductase subunit F	CKS_2080	*nuoF*	0.07	0.04
DSJ_02190	glutamine--fructose-6-phosphate aminotransferase	CKS_1203	*glmS*	0.07	0.04
DSJ_17920	hypothetical protein			0.08	0.03
DSJ_04915	DNA mismatch repair endonuclease MutH	CKS_1357	*mutH*	0.08	0.08
DSJ_16830	NUDIX hydrolase	CKS_2116		0.08	0.05
DSJ_21000	DUF2474 domain-containing protein	CKS_0621		0.08	0.00
DSJ_18685	cysteine--tRNA ligase	CKS_0116	*cysS*	0.08	0.04
DSJ_04200	hypothetical protein	CKS_4935		0.08	0.05
DSJ_02465	glycosyl transferase family 1	CKS_1140		0.08	0.04
DSJ_10820	NUDIX hydrolase			0.08	0.00
DSJ_07475	inositol monophosphatase	CKS_1516	*suhB*	0.08	0.08
DSJ_17805	hypothetical protein	CKS_1925		0.08	0.08
DSJ_09655	aliphatic sulfonates ABC transporter ATP-bindingprotein	CKS_4016	*ssuB*	0.08	0.05
DSJ_02215	phosphate ABC transporter ATP-binding protein	CKS_1197	*pstB*	0.08	0.07
DSJ_02000	type I pantothenate kinase	CKS_0028	*coaA*	0.08	0.08
DSJ_15390	two-component system response regulator	CKS_1847	*cheY*	0.08	0.08
DSJ_21605	GTPase HflX	CKS_0756	*hflX*	0.09	0.05
DSJ_12625	serine/threonine phosphatase	CKS_3338		0.09	0.02
DSJ_16285	CidB/LrgB family autolysis modulator	CKS_2190	*yohK*	0.09	0.08
DSJ_07535	helix-turn-helix domain-containing protein	CKS_1530		0.09	0.06
DSJ_14500	transcriptional regulator CysB	CKS_2790	*cysB*	0.09	0.10
DSJ_01005	Hsp33 family molecular chaperone	CKS_0873	*hslO*	0.09	0.03
DSJ_20685	hypothetical protein	CKS_0546		0.09	0.04
DSJ_18655	AlpA family phage regulatory protein			0.09	0.01
DSJ_14250	YnbE family lipoprotein	CKS_2842	*ynbE*	0.09	0.00
DSJ_04975	exodeoxyribonuclease V subunit alpha	CKS_1343	*recD*	0.09	0.03
DSJ_07665	hypothetical protein	CKS_1736	*ypfN*	0.09	0.07
DSJ_11305	ghrA	CKS_2434		0.09	0.09
DSJ_19540	hypothetical protein			0.09	0.02

DSJ_17420	heme ABC transporter ATP-binding protein CcmA	CKS_2002	*ccmA*	0.10	0.04
DSJ_17145	polyamine ABC transporter permease	CKS_2061		0.10	0.05
DSJ_05460	tyrA	CKS_1240	*tyrA*	0.10	0.06
DSJ_16700	hypothetical protein			0.10	0.10
DSJ_13290	electron transport complex subunit RsxC	CKS_3488	*rsxC*	0.10	0.05
DSJ_02270	hypothetical protein			0.10	0.07
DSJ_08210	replication initiation regulator SeqA	CKS_3764	*seqA*	0.10	0.00
DSJ_19140	thioesterase	CKS_0213		0.10	0.00
DSJ_21965	hypothetical protein	CKS_4466	*yehS*	0.10	0.03
DSJ_13085	oxidoreductase			0.10	0.07
DSJ_06710	ribonuclease HII	CKS_5272	*rnhB*	0.10	0.03
DSJ_07970	GNAT family N-acetyltransferase			0.10	0.06
DSJ_01490	integrase	CKS_0978		0.10	0.09
DSJ_14610	hypothetical protein	CKS_2764	*yciC*	0.10	0.09

[a] 26 genes with 0 read counts (RC) in plant samples (Table 4) were bold; [b] Locus tag and annotation names are from the complete *P. stewartii* DC283 genome (NCBI accession: CP017581); [c] CKS_locus tag and gene names are from the incomplete *P. stewartii* DC283 genome (NCBI accession: NZ_ AHIE00000000.1); [d] ratio of number of RC *in planta* growth over number of RC in LB medium

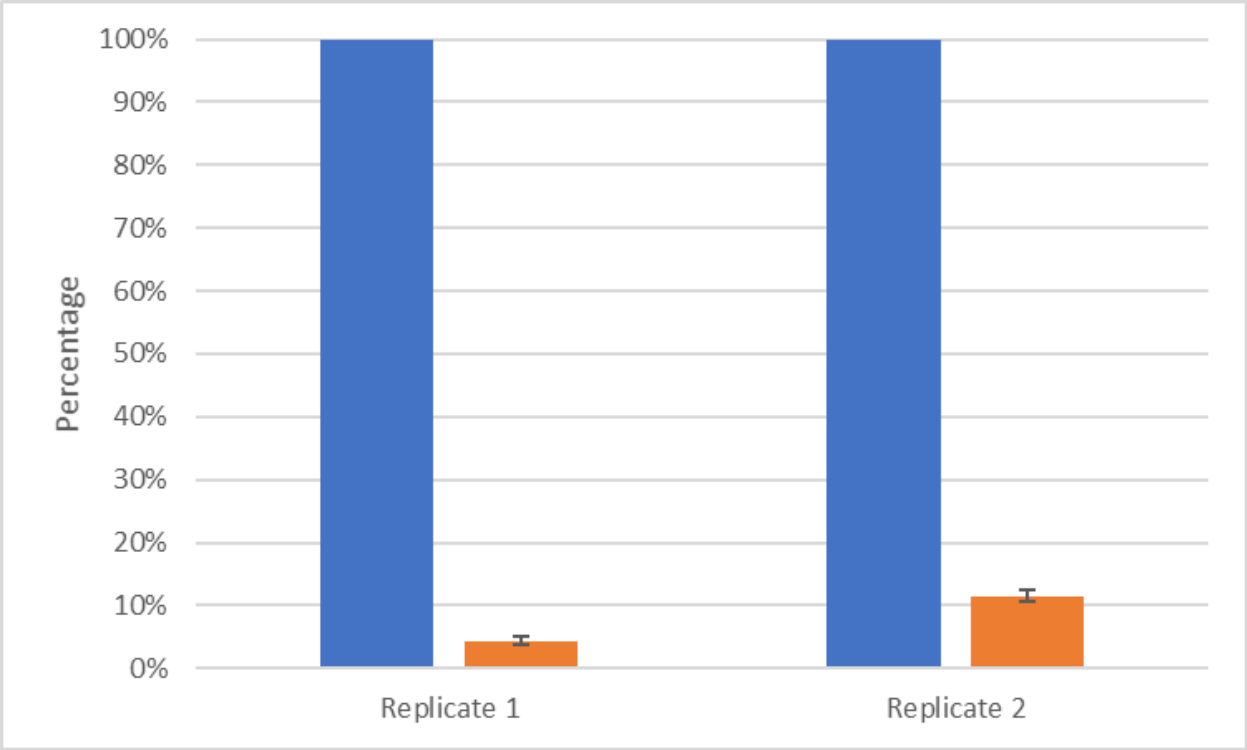

Figure C.1. Estimation of bottleneck effect in the xylem infection model. Blue bars indicate the input CFUs as a percentage of the inoculum prior to infecting the plants, orange bars show the output CFUs that were retrieved from the plant stems after 1 hr post-infection. Error bars in the output samples are standard errors from biological triplicate samples.

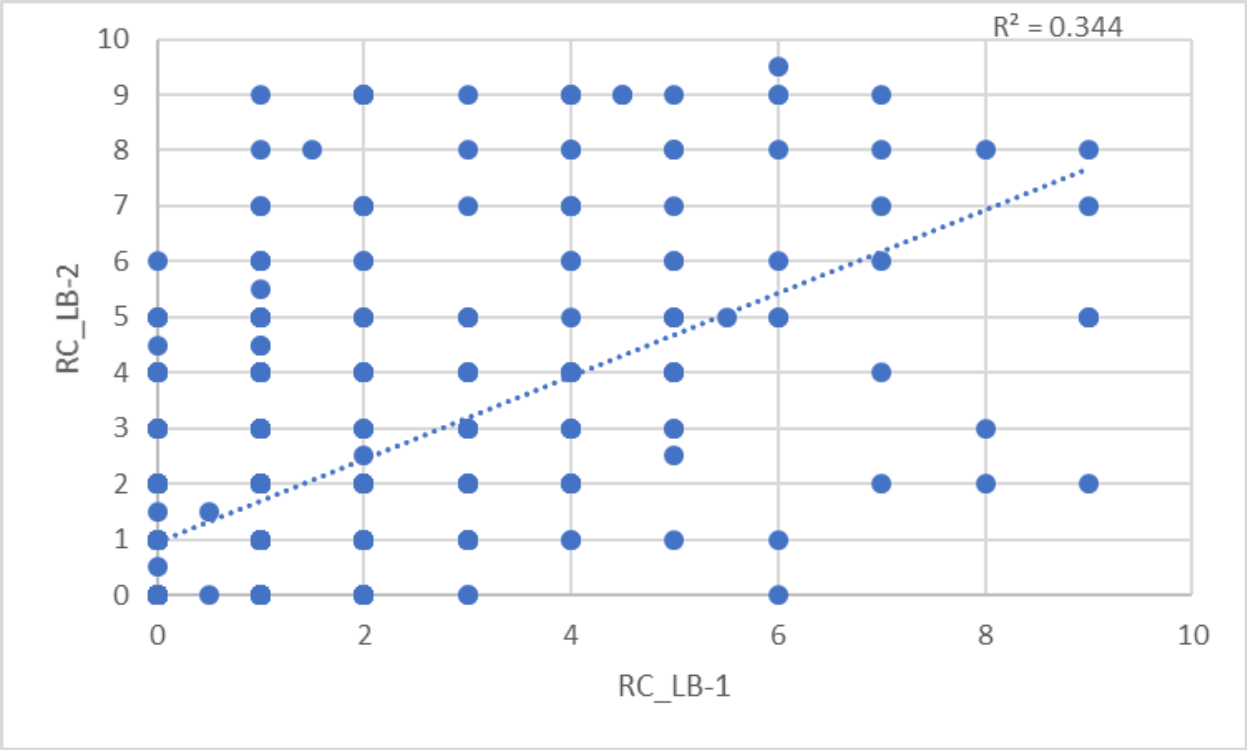

Figure C.2. Low correlation of genes with < 10 read counts (RC) between two LB growth samples. The large difference of the number of mutants in the two libraries grown in LB for these genes might implicate the necessity of these genes in the survival of the bacterium under LB growth condition.

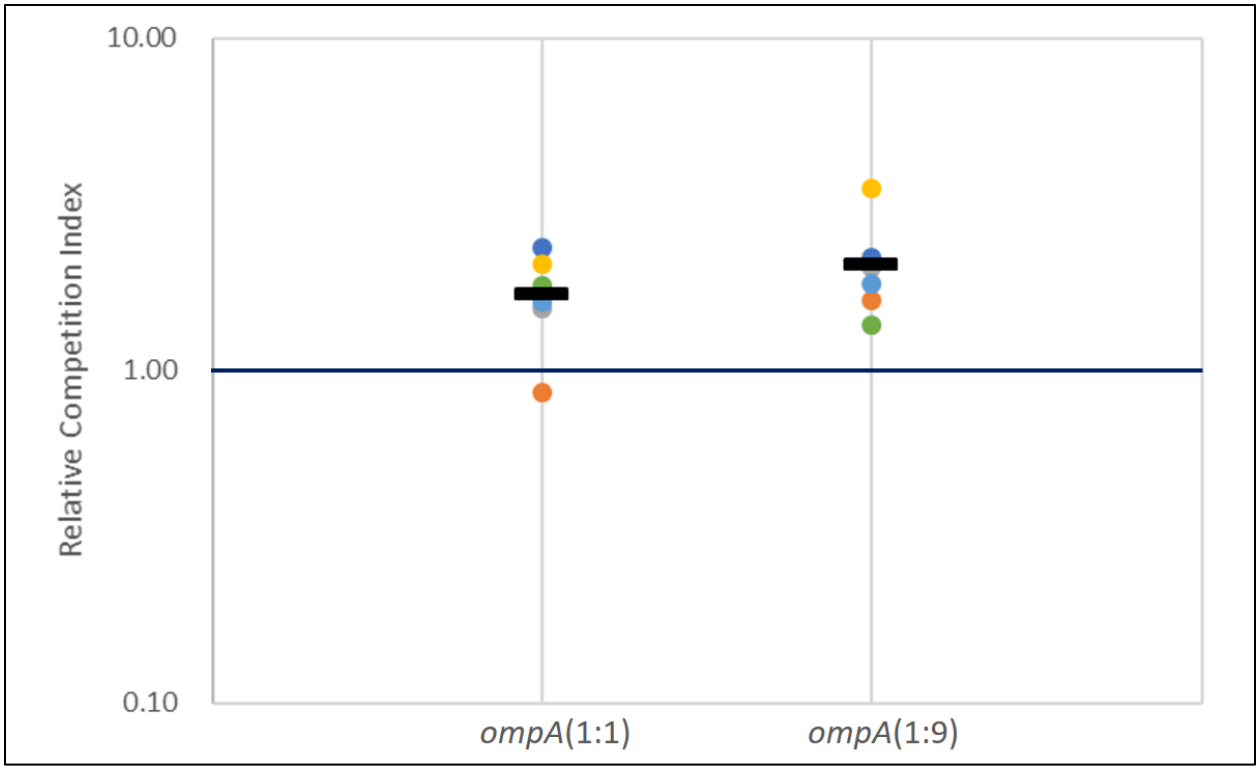

Figure C.3. Repeated results of competition assays for *ompA*.

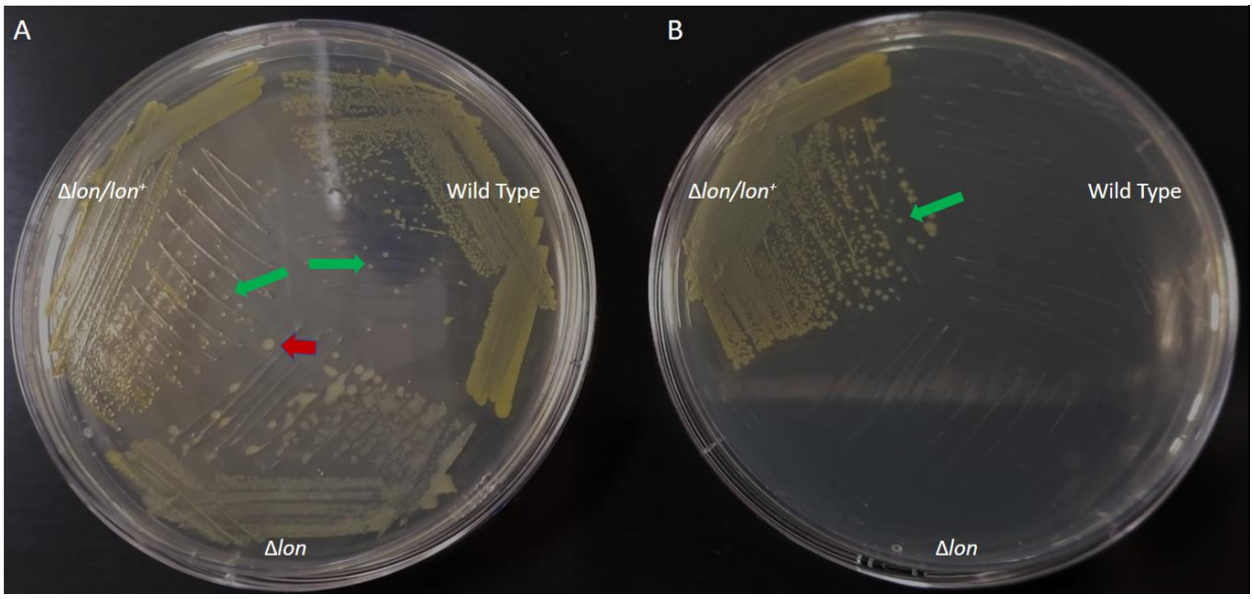

Figure C.4. Colony morphology of the wild-type, Δ*lon*, and Δ*lon/lon*$^+$ strains on LB agar plates. A: LB supplemented with Nal 30 μg/ml; B: LB supplemented with Nal 30 μg/ml and Cm 35 μg/ml. The long arrows (green) point to the regular size colonies of the wild-type and the Δ*lon/lon*$^+$ strains. The short arrow (red) points to a larger size colony from the Δ*lon* strain.

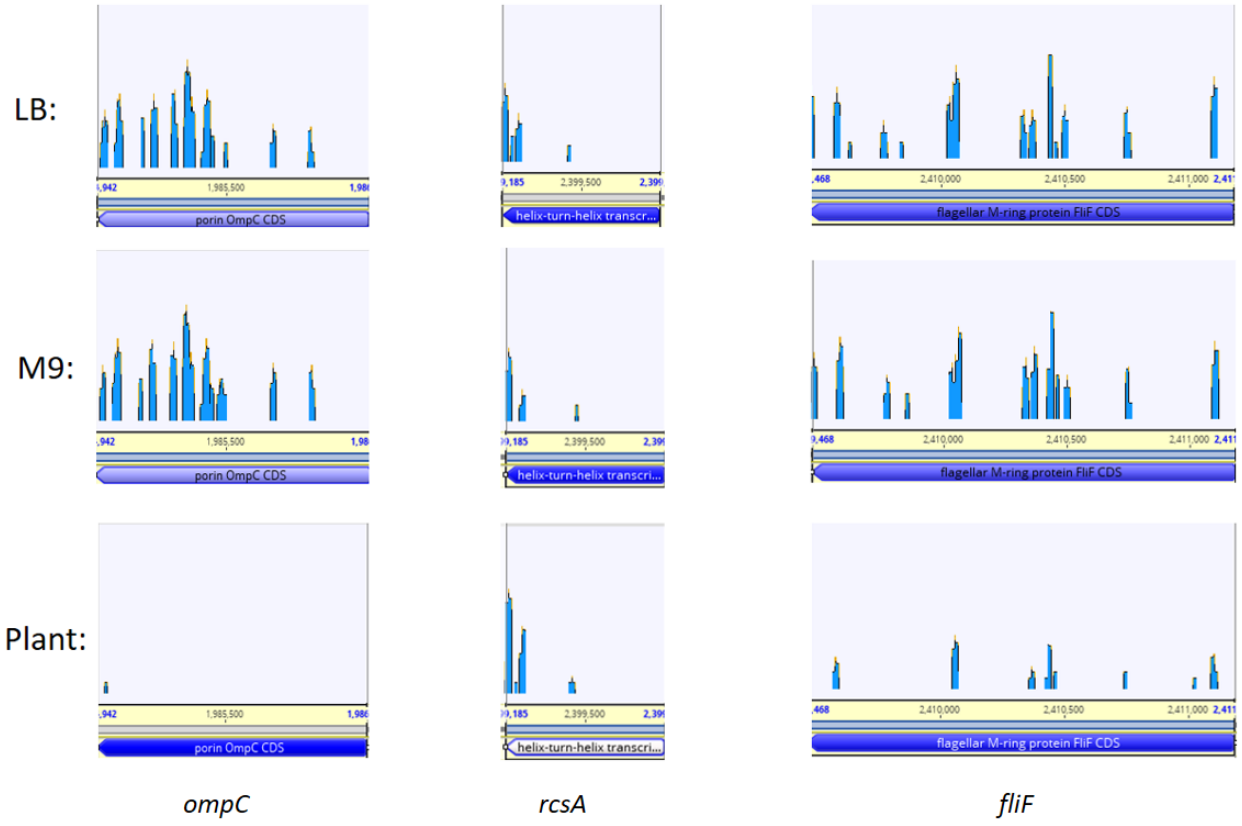

Figure C.5. **Geneious read count plots for three example genes**. (Left) A reduced number of transposons insertions or RC in a gene recovered from the *in planta* samples (e.g. *ompC*) indicates the gene is important in growth/survival. (Center) A higher number of transposon insertion or RC recovered from the *in planta* samples (e.g. *rcsA*) indicates that the wild-type gene actually reduces the bacterial fitness. (Right) Genes with no apparent role *in planta* (e.g. *fliF*) have similar levels (less than ten-fold difference) in transposon insertions or RC across all conditions. See the text for additional details.