Exploring The Virulence Strategy Of The Soft-Rot Plant Pathogen *Pectobacterium carotovorum*

Beth Marie Mole

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Microbiology and Immunology

Chapel Hill
2010

Approved by:

Jeff Dangl

Sarah Grant

Janne Cannon

Robert Bourret

Matthew Wolfgang

Abstract

Beth Mole: Exploring the virulence strategy of the soft-rot plant pathogen *Pectobacterium carotovorum*
(Under the direction of Jeff Dangl and Sarah Grant)

In this dissertation, I explore the virulence strategy of *Pectobacterium carotovorum,* a plant pathogen that causes soft rot disease. *P. carotovorum* infects a wide range of plants, including carrots, beets, and potatoes, and is a significant threat to crop production around the world, causing millions of dollars worth of crop loss each year. In this dissertation, I work with an isolate of *P. carotovorum* from the irrigation pond of a potato farm experiencing a soft rot outbreak in Wisconsin, U.S.A. In ideal conditions, *P. catorovorum* can infect the tubers, stems and leafs of a potato plant. *P. carotovorum* causes disease in part by secreting an array of enzymes that degrade plant cell walls, which causes the characteristic rotting. However, *P. carotovorum* spends a significant part of its life-history in: soil; on invertebrates that feed on crops; and in both surface and ground water, such as the irrigation pond. Additionally, *P. carotovorum* possess a variety of other virulence factors required for plant infection, including a Type III secretion system (T3SS). Type III secretion systems allow a pathogen to deliver virulence factors, called effectors, directly into host cells where they can manipulate plant defense responses. The goal of this dissertation is to better understand how *P. carotovorum* transitions from an environmental microbe to a

pathogen, and the role of the T3SS in that process. In the first chapter, I review how virulence factors, such as plant degrading enzymes and the T3SS, are regulated in *P. carotovorum,* and compare its regulation strategy to that of other plant pathogens. In the second chapter, I present an analysis, published with our collaborators, of the genome sequence of *P. carotovorum,* which we draw upon in the subsequent chapters. The third chapter evaluates transcriptional data during tuber, stem and leaf infections and suggests a tissue-specific role for the T3SS during infection. Lastly, the fourth chapter presents data that suggests carbon source availability is involved in tissue-specific virulence regulation. In conclusion, this dissertation furthers our understanding of the virulence strategy of *P. carotovorum.*

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: Global virulence regulation networks in *Pectobacterium carotovorum* and model plant pathogens

ABSTRACT:

This chapter provides an introduction and framework to the dissertation project; relevant background of the soft-rot pathogen *Pectobacterium carotovorum* is covered and followed by a comprehensive review of phytopathogen virulence strategies. The overall goal of this dissertation is to further our understanding of the virulence regulation and strategy of *Pectobacterium carotovorum* in the context of environmental signals and life-history. *P. carotovorum* is an enterobacteriacea family member and ubiquitous plant pathogen that regulates virulence determinants using a complex set of regulators. Generally, phytopathogens coordinate multifaceted life histories and deploy stratified virulence determinants via complex global regulation networks. The review dissects the global regulation of four distantly related model phytopathogens to evaluate large-scale events and mechanisms that determine successful pathogenesis. Overarching themes include dependence on centralized cell-to-cell communication systems, pervasive two-component signal transduction systems, post-transcriptional regulation systems, AraC-like regulators, and sigma factors. Although these common regulatory systems control virulence, each functions in different capacities, and to differing ends, in the diverse species. Hence, the virulence regulation network of each species determines its survival and success in various life-histories and niches. This review will highlight unanswered questions about the virulence regulation and strategy of *P. carotovorum,* which will be addressed in the following chapters.

INTRODUCTION:

***Pectobacterium carotovorum* is an economically significant plant pathogen and a member of the Enterobacteriacea family.** The enterobacteriacea family encompasses a wide range of bacterial pathogens that infect plants, animals, and fungi, and can be found in soil, water and atmospheric samples. Members of this family have had dramatic impacts on human history, such as *Yersinia pestis,* which caused the black plague [1], and *Vibrio cholera,* which causes cholera, a disease still endemic in many regions [2]. Many enterobacteriacea family members were sequenced and well studied before the onset of this work, and therefore the family is ideal for comparative bacterial genetic and environmental studies. *Pectobacterium carotovorum,* a plant pathogen in the family, is a further ideal candidate because it is ubiquitous in soil, water, and atmospheric samples worldwide and can infect species from 24 orders of plants (reviewed in [3]). *P. catorovorum* can subsist in an environment until ideal conditions arise, such as plant damage from hail storms, hot weather and/or high humidity, and can cause widespread rot during the growing season or crop storage. *P. carotovorum* is a significant threat to crop production in the U.S. and is thought to cause millions of dollars worth of crop destruction each year. In addition to an agriculturally relevant study system, *P. carotovorum* demonstrates significant genetic variation, even among isolates from clonal hosts, such as clonally propogated potatoes [4], making it an ideal candidate for biogeographical analysis and comparative genetic studies of strains isolated from different hosts. The broad goals of this

dissertation are to understand virulence determinants and regulation of *P. carotovorum* in the context of environmental signals and related pathogens.

***P. carotovorum* is a brute-force and stealth pathogen.** *P. carotovorum* uses a variety of virulence factors to elicit disease on a host, which include a type II secretion system (T2SS) [5-7], adhesions [8], metalloproteases [9], and a type III secretion system (T3SS) [10]. *P. carotovorum* was first described as a brute-force pathogen and defined as a necrotroph because it secretes a series of plant-cell-wall-degrading enzymes (PCWDE) to macerate host cells, which cause characteristic rot and ultimately death to the host. PCWDEs are secreted through a T2SS, which is a specialized outer-membrane transport system dependent on generalized secretion, Sec- or TAT-dependent systems. The T2SS is responsible for secreting pectate lyases, polygalacturonases, and pectinases, which degrade pectin, a principle component of the plant cell walls.

***P. carotovorum* also possess a T3SS, which delivers virulence factors, referred to as effectors, directly into the cytoplasm of a neighboring host cell**. Many enterobacteriacea family members, and Gram negative pathogens in general, require a T3SS for virulence [11]. Delivered effectors often manipulate plant hosts by repressing defense responses and altering nutrient availability [12], which allows the bacteria to stealthly persist in the host. Plants induce innate defense responses, such as thickening the cell wall with callose and secreting reactive oxygen species, when they detect microbial-associated

molecular patterns (MAMPs), like LPS and flagellin (reviewed in [13]). T3SS help the invading bacteria evade detection and starvation by delivering effectors that inhibit these plant defenses. Plants, in turn, have adapted to detect delivered effectors in the cytoplasm as a means for detecting invading bacteria and reinstating defense responses. Proteins, coined recognition (R) proteins, directly or indirectly detect a bacterial effector and can subsequently elicit programmed cell death, also known as Hypersensitive Response (HR). This defense response can be seen as a dead patch on the leaf at the site of infection and results in inhibition of bacterial growth. Since there is evolutionary pressure on the pathogen to deliver unrecognized-effectors, and on the plant to continually recognize bacterial effectors, bacterial pathogens and their plant hosts are locked in an evolutionary arms race. Many well-studied plant pathogens require a T3SS for virulence and large suites of effectors to deliver to host cells. *P. syringae* species, for example, have been found to encode upwards of 30 effector proteins that are delivered via a T3SS [14]. Different *P. syringae* isolates contain different collections of effectors and effectors can be easily acquired by horizontal transmission or lost when pressured.

The T3SS of *P. carotovorum* is most closely related to that of *P. syringae.* The cluster of genes that encode the secretion apparatus is regulated by an alternative sigma factor, HrpL, the homolog of which plays the same role in *P. syringae* [15]. Only one effector has been identified in *P. carotovorum*, a homolog of AvrE in *P. syringae*, DspE. The number and variation in *P. carotovorum* effectors is, therefore, still unknown [16]. However, the possession

and requirement for a T3SS during pathogenesis suggests that other effectors exist. In the following chapters, we specifically address the role of the T3SS in *P. carotovorum* virulence and identify T3SS delivered effectors.

Virulence determinants, such as the T2SS and the T3SS, in *P. carotovorum* are regulated by a complex global regulation network. An over-arching question in our understanding of the pathogenesis of *P. carotovorum* is how this pathogen transitions from an environmental microbe to a plant pathogen of various hosts and then coordinately deploys stealth virulence determinants, like the T3SS, and brute-force, the T2SS. Much work has already begun to shed light on the wealth of regulators, which *P. carotovorum* possesses (reviewed in [7, 17-21]. The following section of this chapter will review virulence regulatory networks of *P. carotovorum* in relation to other plant pathogens as a foundation for this dissertation project. Chapter three provides comprehensive transcriptional profiling of *P. carotovorum* during multiple infection routes, which reveals a requirement of the T3SS during leaf infections, but not other infection routes. Chapter four covers novel work that addresses how metabolic cues act as key environmental signals in virulence deployment. The data in this dissertation furthers our understanding of the virulence strategy of *P. carotovorum* by identifying metabolic cues and differential virulence expression among infection routes. To understand how this novel data furthers the field, we must first review what we know about virulence regulation and strategies of plant pathogens.

GLOBAL VIRULENCE REGULATION IN PLANT PATHOGENS:

Global virulence regulation networks dictate life-histories and infection strategies of plant pathogens. Phytopathogens survive in diverse environments, not only as pathogens but also as benign epiphytes on plant surfaces or saprophytes in soil and water. Consequently, expression of virulence factors and behaviors associated with virulence must be coordinated for energy conservation, appropriate disease development, evasion of host defense, and eventual dispersal. A phytopathogen's survival, therefore, relies on a controlled global virulence regulation network. In this introductory chapter, we discuss such networks in four distantly related and well studied phytopathogens (Figure 1). *Ralstonia solanacearum* (hereafter *R.s.*) is a hemibiotrophic agent (with initial biotrophy followed by necrotrophy during infection) of bacterial wilt diseases across a wide host-range (reviewed in [22]). *R.s.* also exists as a saprophyte in soil until it invades the roots of susceptible plants. *R.s.* colonizes xylem tissue and migrates to aerial parts of the plant where it accumulates and disables the plant's vascular system (Figure 1a). *P. carotovorum* (hereafter *P.c.*) is a necrotrophic soft rot pathogen with a large host-range (reviewed in [17]). *P.c.* can live as an epiphyte or as a saprophyte in soil and ground water until it encounters a susceptible host (Figure 1b). *Xanthomonas campestris* pv. *campestris* (hereafter *X.c.c.*) is a hemibiotrophic, narrow host-range agent of

black rot (reviewed in [23]) (Figure 1c). *Pseudomonas syringae* (hereafter *P.s.*) is the hemibiotrophic agent of bacterial speck, spot and rot diseases typically associated with narrow host range (reviewed in [12]) (Figure 1d). Both *P.s.* and *X.c.c.* exist as epiphytes on plant surfaces until opportunity allows them to enter the intercellular apoplastic space inside leaf tissue. These disease

Figure 1.

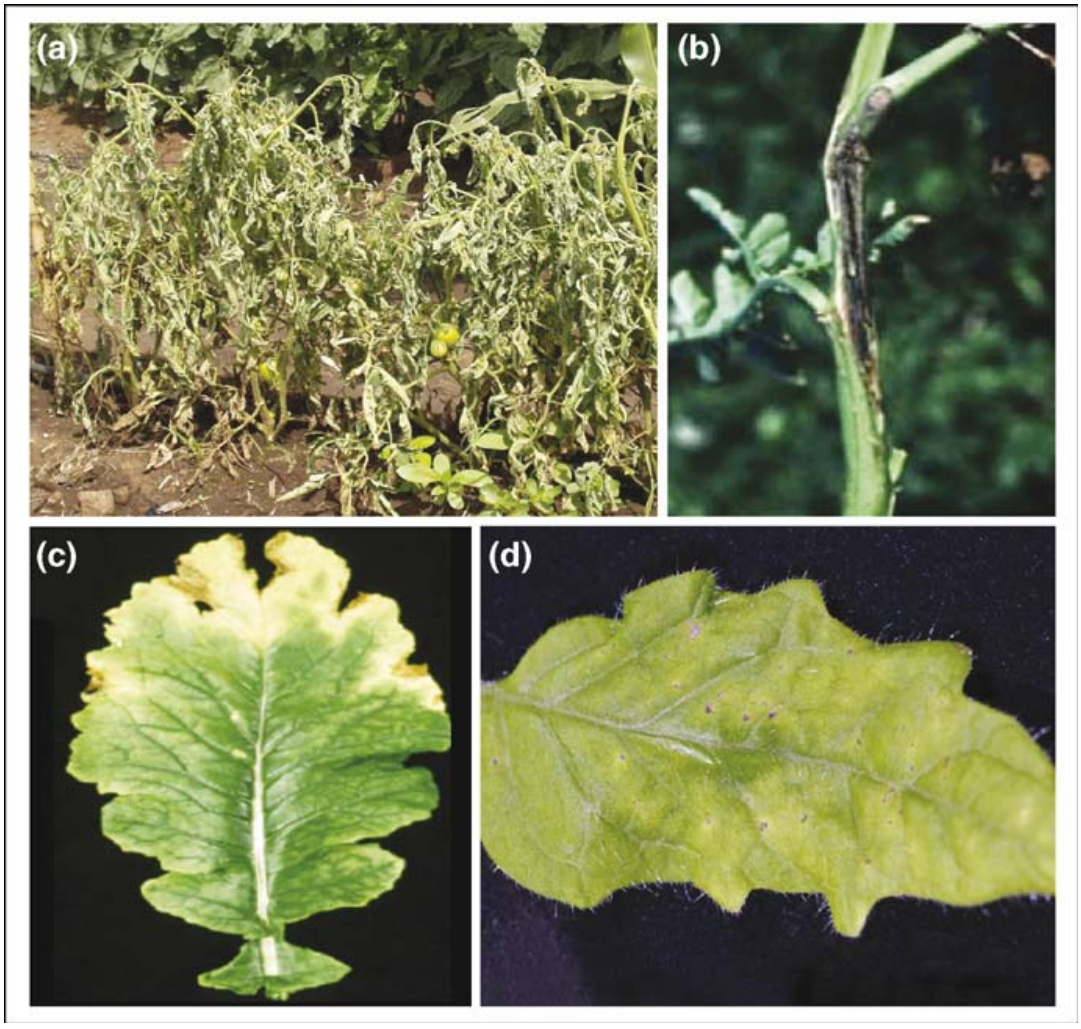


Figure 1. Disease symptoms of selected phytopathogens. a. Bacterial wilt of *Ralstonia solanacearum* on tomato (courtesy of Caitilyn Allen, University of Wisconsin-Madison). **b.** Stem rot by *Pectobacterium carotovorum* subsp.

carotovorum on potato (courtesy of Amy Charkowski, University of Wisconsin-Madison). **c.** Black rot by *Xanthomonas campestris* on cabbage (courtesy of Max Dow, University of Ireland). **d** *Pseudomonas syringae* bacterial spot on tomato (courtesy of Marc Nishimura, University of North Carolina-Chapel Hill).

---

stages and life-history transitions require information sensed from the environment, the host, and the pathogen population. These inputs are received and integrated by central regulators to produce survival and virulence outputs through a global virulence regulation network.

The mechanisms that control expression of vital T3SS, toxins, extracellular enzymes, and various behaviors are discussed below. This section begins by presenting cell-to-cell communication systems and their influence on prominent regulators in each species. Then the chapter will compare two-component regulators that report information about the environment to central regulator(s) or influence distinct aspects of virulence independently. The chapter will finish with a discussion of common transcription regulators that control T3SS expression, but have global impacts as well. By comparing global virulence regulation strategies of these four pathogen groups we can understand common mechanisms essential to plant infection and areas of divergence that enable each pathogen to take advantage of its biological niche. This chapter will also help identify areas of *P. c* regulation and pathogenesis that will be addressed in the coming chapters.

I. Cell-density-dependent regulation of virulence factors is crucial for all four phytopathogens.

Cell-to-cell communication systems enable temporally coordinated gene expression within bacterial populations. The infection strategies of phytopathogens, which often require swift global changes in gene expression and physiology in response to environmental cues, are particularly reliant on cell-to-cell communication to coordinate critical steps in pathogenesis. Some of the communication systems of phytopathogens are paradigmatic quorum sensing systems, analogous in mechanism to the conserved LuxR/I system of *Photobacterium fischeri*, while others represent mechanistically distinct systems particular to related pathogen groups. Here, we compare the main features and functions of cell-to-cell communication systems with respect to central regulators in each of our model species.

Cell-to-cell communication in *R.s.* is responsible for the transition from early to late stage infection

R. s. uses a genus-specific communication system that senses the diffusible 3-OH-palmitic acid methyl ester (3-OH-PAME) signal. The 3-OH-PAME system allows *R.s.* to make the transition from early to late stage pathogenesis by controlling the activity of the global virulence regulator, PhcA (Figure 2a). The 3-OH-PAME system is encoded by the *phc* (phenotypic conversion) operon [24] which encodes the 3-OH-PAME synthase, PhcB, the membrane-bound sensor, PhcS, and a downstream response regulator, PhcR [25]. At low cell-density, or in

unconfined conditions, the global virulence regulator, PhcA, is inhibited by the 3-OH-PAME system. When the bacterial population density reaches ~10^7 cells/mL and/or the concentration of 3-OH-PAME reaches 5nM, 3-OH-PAME binds to the membrane-bound sensor PhcS [26] and PhcA is released from repression by PhcS. PhcA then either directly or indirectly activates virulence factors required for late-stage infection, including exopolysaccharides (EPS) [27], and cellulase [24] and represses virulence factors required for early-stage infection including the T3SS [28], swimming motility and siderophore expression [29].

The 3-OH-PAME system also up-regulates an HSL-dependent quorum sensing system via PhcA activity. The quorum sensing system is encoded by *solR* and *solI*, which are *luxR* and *luxI* homologues, respectively [30]. SolR activates expression of at least one gene, *aidA.* The function of the novel AidA protein is unknown, and other SolR-regulated genes have yet to be identified. *solR* and *solI* are members of the 3-OH-PAME communication system regulon which controls the phenotypic switch from a saprophyte and early stage colonizer to a full blown pathogen. Because co-regulated genes generally function similarly, the SolR/SolI system might be expected to play a role in virulence. Surprisingly, *solI* and *solR* mutants have been shown to retain full virulence when inoculated into cut petioles [30]; however reduction in the virulence of these mutants might be evident if they were tested using different infection conditions such as soil soaking, which more closely mimic a natural infection.

Cell-to-cell communication in *P.c.* regulates global virulence expression, including T3SS and extracellular enzymes, by controlling the expression of the regulator, RsmA

Just as in *R.s.,* the cell-to-cell communication systems of *P.c.* are central to virulence regulation. However, the communication systems are more complex in their regulatory mechanisms (and, sadly, their nomenclature). The *P.c.* cell-to-cell communication systems are responsible for regulating the T3SS, PCWDE, and antibiotic production. *P.c.* employs quorum sensing systems that include up to three transcription activators which are responsive to two AHL molecules that, in turn, are encoded by one synthase (Figure 2b). The LuxI synthase homologue, ExpI (also referred to as CarI, AhlI, HslI), was first described by Pirhonene *et al.* [31] as the exoenzyme production inducer. ExpI can synthesize 3-oxo-C6-HSL or 3-oxo-C8-HSL. Strains that produce both variations of AHL have been designated class I strains, while strains that produce only 3-oxo-C6-HSL have been designated class II [17]. Once AHL signaling molecules accumulate, they can interact with CarR, ExpR1 or ExpR2. The most straightforward of these is the CarR regulator. CarR binds 3-oxo-C6-HSL and subsequently binds the *carA* promoter, which controls the *car* operon that encodes the carbapenem antibiotic [32] (Figure 2b). The *car* operon is also controlled by the transcription regulator, Hor, but by unknown means. The other two LuxR homologues, ExpR1 (also known as ExpR and EccR) and ExpR2 (also known as VirR), directly inhibit virulence in the absence of threshold levels of AHL by up-regulating *rsmA* [33]. RsmA is a member of the post-transcriptional Rsm system and acts to destabilize

mRNA transcripts that encode PCWDEs, including cellulase, pectate lyase, and protease. The Rsm system also includes, *rsmB,* which binds to, and down regulates RsmA, allowing translation of RsmA-targeted mRNAs. Threshold levels of 3-oxo-C8-HSL bind to ExpR1 while 3-oxo-C6-HSL binds to ExpR2. Binding of their cognate AHL inactivates ExpR1 and ExpR2, inhibiting *rsmA* expression and freeing the mRNA transcripts encoding PCWDEs [33, 34]. Mutant *P.c.* lacking *expI* are unable to produce AHLs and PCWDEs and are consequently avirulent [31]. Thus, *P.c.* quorum sensing feeds into the central coordination of virulence for successful infection via novel, synergistic negative regulation by two LuxR homologues ExpR1 and ExpR2.

Post-transcriptional Rsm-mediated regulation is significant in both P.c. and P.s.

The Rsm system plays a significant role in the virulence regulation of *P.c.* but it is broadly distributed throughout the eubacteria, and thus might be a common central player in virulence regulation [35]. The Rsm system is made up of the proteins RsmA, and RsmC and the regulatory RNA *rsmB* [34]. RsmA is the main regulator and targets mRNA transcripts in order to regulate virulence as well as cellular functions. RsmA and the related CsrA from *E. coli* are thought to act similarly; they bind mRNA transcripts using a KH domain and can either stabilize or promote the decay of targeted transcripts [36]. CsrA destabilizes messages by binding near the ribosome binding site, thereby interfering with translation and leading to transcript degradation by RNases [37]. In *P.c.,* RsmA is independently

down-regulated by the quorum sensing system and the stationary sigma factor, RpoS [38] (Figure 2b). RsmA, in turn, targets and destabilizes the transcripts of PCWDEs which are crucial virulence factors [39] and HrpL, the alternative sigma factor that regulates expression of the T3SS (reviewed in [17]) (Figure 2b)

Figure 2.

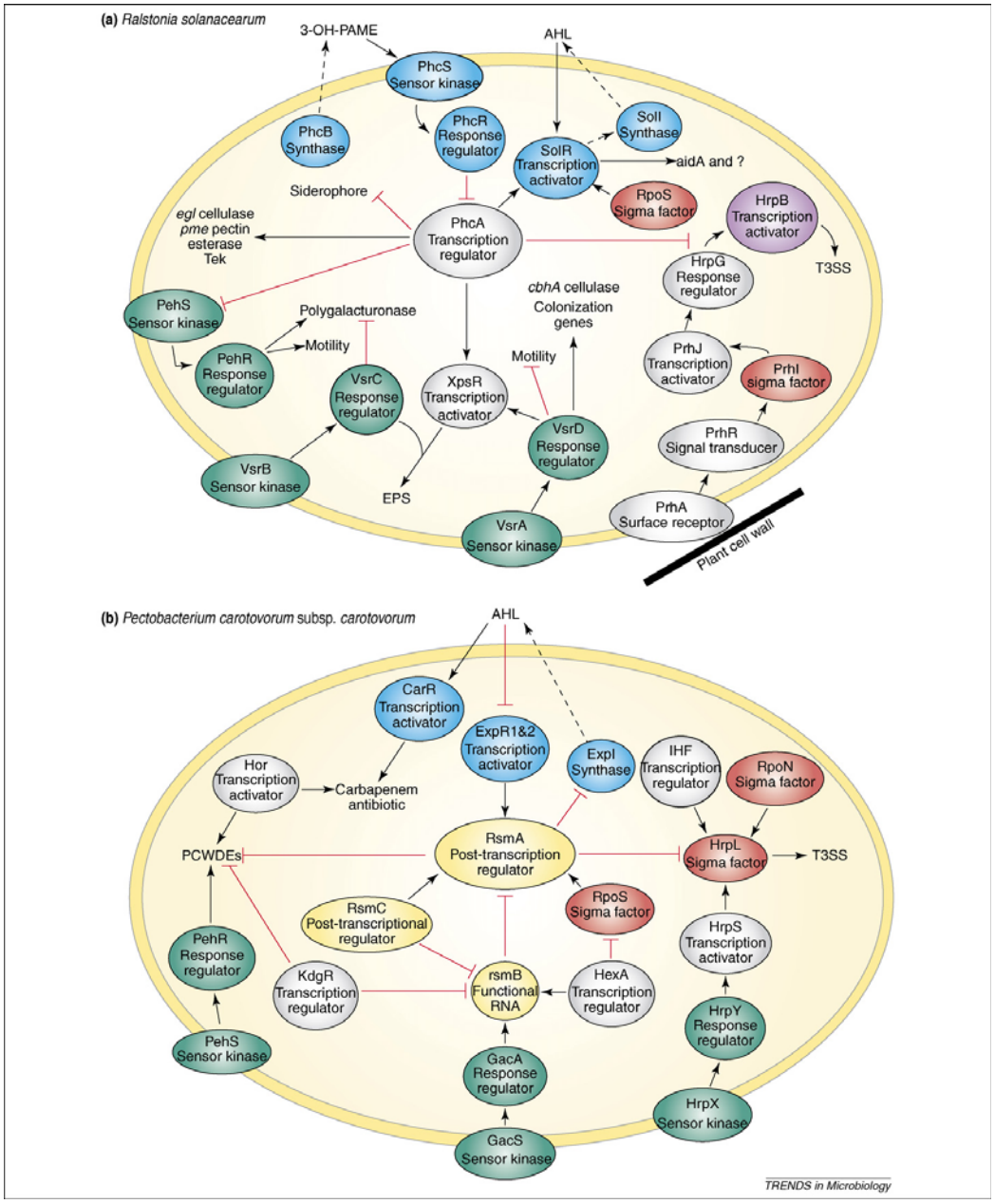


Figure 2. The global virulence regulation of *R.s.* and *P.c.* Items in: blue are associated with cell-to-cell communication; green are TCST systems; red are sigma factors; yellow are part of the Rsm system; and purple are AraC-like transcription regulators. Cell-to-cell communication systems are located at the top of each figure, TCST systems are located on the bottom, extracellular enzymes and EPS regulation are found on the left, T3SS regulation is found on the right, and global regulators are situated in the center. Arrows represent positive regulation, barred arrows represent negative regulation, dashed arrows represent synthesis, and open arrows represent unknown interactions. **a.** *Ralstonia solanacearum*. PhcA is a global virulence regulator in *R.s.* controlling motility, polygalacturonase, EPS, and the T3SS. PhcA is controlled by a species specific cell-to-cell communication system and, in turn, PhcA regulates the secondary quorum sensing system. **b.** *Pectobacterium carotovorum* subsp. *carotovorum*. The prominent virulence mechanism in *P.c.* is the post-transcriptional Rsm system. The Rsm system is mediated by cell-to-cell communication and the GacA/S TCST system and controls the T3SS and cell wall degrading enzymes.

---

rsmB (also known as *aepH* in *P.c..*) and *csrB* in *E. coli,* are non-coding, functional RNAs that act as antagonists to RsmA and CsrA (reviewed in [40]). *rsmB* is up-regulated by the GacA/S TCST system [41] and HexA, a global regulator that also inhibits RpoS expression [20]. *rsmB* is repressed by the KdgR regulator which is responsive to breakdown products of plant cell walls, such as

2-keto-3-deoxygluconate, KDG [42]. The *rsmB* structure contains multiple stem loops with GGA motifs thought to be binding sites of RsmA [43]. CsrA and *csrB* can be co-purified in a stoichiometry of 18:1, suggesting that each *csrB* molecule binds 18 CsrA molecules and thus efficiently inhibits its binding to target transcripts [19]. In *P.c. rsmB* also efficiently inhibits RsmA. *rsmB* exists in two forms, a 479 nucleotide, unprocessed form and a processed 259 nucleotide form which is thought to sequester RsmA via nine CsrA-like binding motifs [19]. Lastly, RsmC, also known as HexY, negatively regulates *rsmB* and, directly or indirectly, activates *rsmA* expression [18]. RsmC has no known homologues outside of *P.c.* and has no known DNA binding motif. The complete regulon of RsmA and the Rsm regulation system is unknown.

The cell-to-cell communication in *X.c.c.* is crucial to global virulence regulation and basic cellular functions and employs cyclic di-GMP as a second messenger

X.c.c. posseses a previously unknown cell-to-cell communication system that might be shared among bacterial species [44]. This system relies on a diffusible signal factor (DSF) that has been linked to the regulation of motility, toxin and oxidative stress resistance, aerobic respiration [45], biofilm dispersal, and extracellular enzyme and EPS production [46] (Figure 3a). DSF (cis-11-methyl-2-dodecenoic acid) is synthesized by products of the *rpf* virulence regulation cluster genes *rpfB* and *rpfF* [44]. DSF accumulates in early stationary phase and is

sensed by a unique two-component signal transduction system consisting of RpfC and RpfG. This leads to degradation of cyclic di-GMP [47]. Cyclic di-GMP

Figure 3.

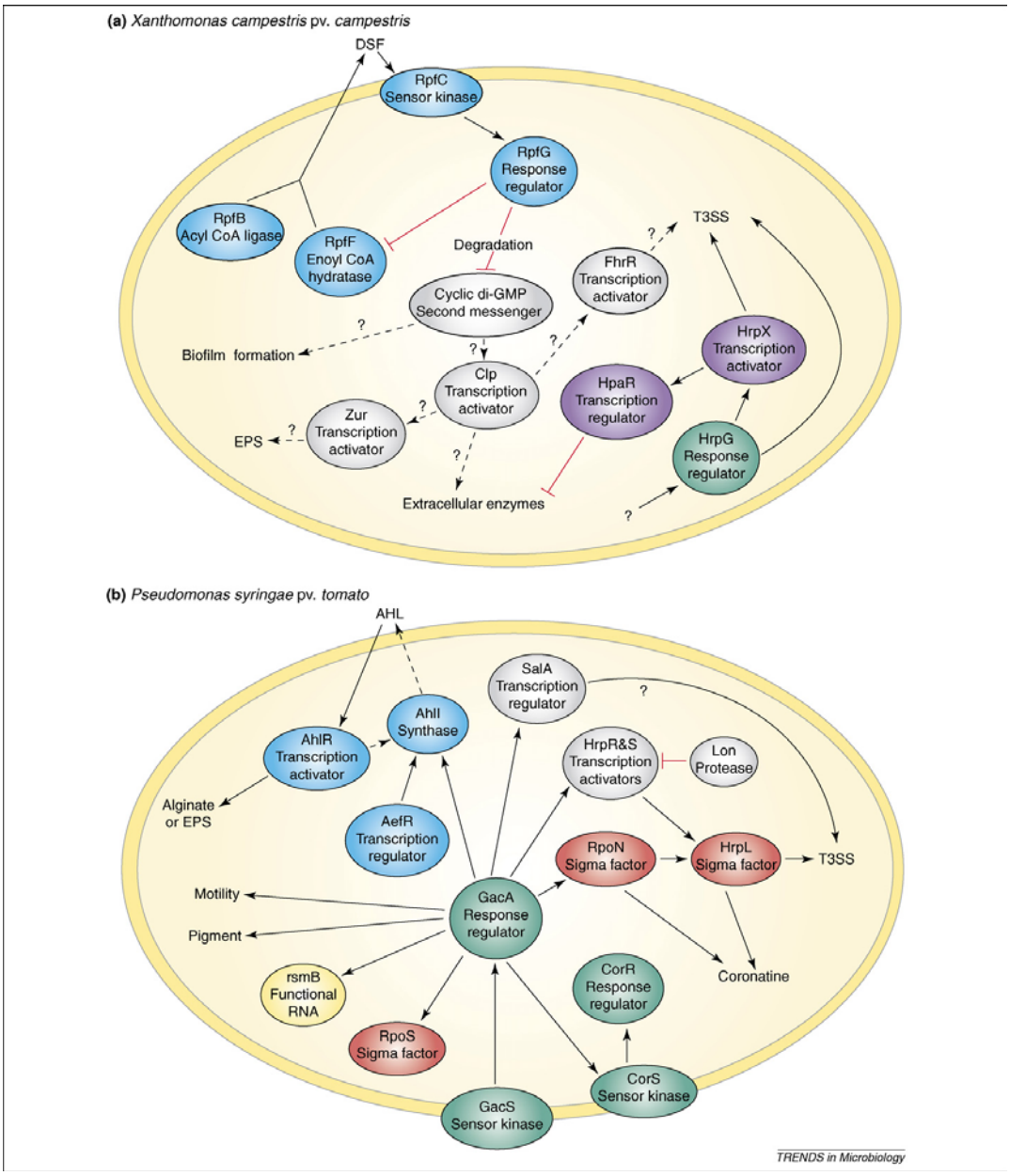


Figure 3. The global virulence regulation of *X.c.c.* and *P.s.* The coloration, layout and arrow designations are the same as Figure 2. **a**. *Xanthomonas campestris* pv. *campstris*. The global regulator is the small molecule cyclic di-

GMP. Cyclic di-GMP is upregulated by cell-to-cell communication via a unique TCST system. Cyclic di-GMP levels subsequently affect biofilm formation, extracellular enzymes, EPS and the T3SS. **b.** *Pseudomonas syringae* pv. *tomato*. The global regulation system in *P.s.t.* is the GacA/S TCST system. GacA is known to directly or indirectly regulate quorum sensing, coronatine production, the T3SS, motility, and pigment production.

---

was first described as an activator of bacterial cellulose synthase by Ross *et al.* [48] but was subsequently ascribed global regulatory functions (reviewed in [49]). Decreased levels of cyclic di-GMP in *X.c.c.* lead to dispersal of biofilms, increased production of EPS and extracellular enzymes, and alterations in T3SS expression [50]. Cyclic di-GMP is also involved in the regulation of iron up-take, metabolic activities, LPS production, multi-drug resistance and detoxification, but the regulatory network(s) responsible for cyclic di-GMP–dependent activation are poorly understood. He *et al.* used microarray data to identify three heirarchal transcription regulators (Clp, Zur, FhrR) downstream of cyclic di-GMP activation that are involved in EPS, T3SS, and extracellular enzyme production among other cellular functions [51] [52]. This novel and pleiotropic regulation system links virulence to basic, essential cellular functions.

***P.s.* cell-to-cell communication impinges on a central regulation system, GacA/S, and is needed for both early and late stages of disease**

Unlike the three systems described above, the cell-to-cell communication system of *P.s.* does not appear to be the first step in the regulatory hierarchy of virulence, controlling a regulatory molecule that subsequently integrates many virulence responses. Instead, it is the virulence integrating GacA/S TCST system (described below) that controls the AhlI/R (LuxI/R-homologous) quorum sensing system in combination with the TetR-family transcription activator AefR [53] (Figure 3b). The AhlI/R (LuxI/R-homologues) system is significant for the regulation of intercellular host tissue maceration [54] and epiphytic fitness [55] because it controls EPS production and swarming motility in diverse environments [54]. By regulating swarming motility, quorum sensing becomes a means to disperse epiphytic aggregates on the leaf surface. Quorum sensing mutants are also unable to macerate tissue at later stages of infection. Therefore, the AhlI/R system is significant for regulating virulence at various stages of infection. The complete regulon controlled by AhlI/R, and the mechanisms by which it is regulated, are unknown.

II. Two-Component Signal Transduction (TCST) systems provide environmental signal input to global virulence regulation

TCST systems are commonly used by bacteria to sense and adapt to extracellular environmental signals. The two common, and very well studied, components are a membrane-bound sensor kinase and a cognate, cytoplasmic response regulator (reviewed in [56]). Subsequent to signal perception, the cytoplasmic response regulator is either activated or inactivated by

phosphorylation. The response regulator can carry out a variety of tasks, including transcriptional regulation and protein-protein interactions. In the phytopathogens discussed in this review, TCST systems are typically used to directly regulate single or a small number of related virulence determinants, while only a few TCST systems have global effects on virulence.

TCST systems in *R.s.*, *P.c.* and *X.c.c* complement global virulence regulation

In *R. s.*, it appears that global virulence regulation is coordinated by quorum sensing (see above), and TCST systems play ancillary roles in virulence regulation. For example, the sensor kinase / response regulator pairs VsrA/D and VsrB/C, respectively, both act in combination with PhcA (discussed above) to regulate a key virulence factor, EPS (reviewed in [22]), Together, activated VsrD, and PhcA activate XpsR, an atypical response regulator [27, 57] (Figure 2a). Activated XpsR, together with activated VsrC initiates transcription of the 16kb *eps* operon, responsible for production exopolysaccharide 1 (EPS I) [22, 58]. Additionally, *R.s.* uses the NtrB/C-like TCST system PehS/R. PehS/R acts in opposition with VsrC to activate the transcription of *pglA,* one of three polygalacturonases used to degrade the plant cell wall. In addition VscR activates swimming and twitching motility (Figure 2a) [59].

Similarly, *P.c.* deploys TCST systems that are activated independently from the global regulator RsmA and play ancillary roles in virulence regulation. For

example, the PehS/R TCST system in *P.c.* also regulates PCWDEs (Figure 2b) as the system with the same name does in *R.s.* However the PehS/R in *P.c.* is not a member of the NtrB/C family like the *R.s.* PehS/R. Instead, it is a member of the EnvZ/OmpR family of the TCST systems [60]. Notably, this system is homologous to the PhoP/Q system that regulates virulence throughout the Enterobacteria, suggesting that the theme of ancillary roles for TCST systems in virulence regulation extends beyond phytopathogens [60].

The OmpR-family response regulator of *Xanthomonas campestris* species, HrpG, regulates the T3SS [61] (Figure 3a). Although a cognate sensor kinase for this response regulator has not been identified, this finding suggests that a TCST might be involved in T3SS regulation, in addition to the cyclic di-GMP interactions, which are regulated by quorum sensing (discussed above). The lack of an identified cognate sensor kinase suggests the possibility of cross-talk between multiple systems and/or redundancy in the function of a sensor kinase. Future studies on HrpG activation will be useful in understanding this seemingly incomplete TCST system, and will define the cues for T3SS system expression.

These are exemplary cases of TCST systems that contribute supplemental regulatory function to global virulence regulation networks. However, a recent study using reporter gene studies has shown *R.s.* VsrA is a direct regulator of motility, T3SS and quorum sensing, suggesting more than a supplemental role in virulence regulation that may involve secondary metabolism [62] (Figure 2a).

Although the mechanisms of this regulation are unknown, further studies and fully identified regulons of TCST systems might shift our view of their roles in global virulence regulation.

The GacA/S and RpfC/G TCST systems are exceptional in their global impact on virulence

The GacA/S system is crucial to virulence regulation in both P.c. *and* P.s.

In contrast to all other TCST systems discussed so far, the GacA/S TCST system plays a key global role in virulence regulation. The Gac system is found in a variety of Gram-negative bacteria and its name was coined from *Pseudomonas fluorescens* as the global regulator of antibiotic and cyanide production [63], although the homologue LemA (lesion manifestation) was identified first in *P.s.* [64]. The Gac system in *P.c.* and *P.s* interacts with the Rsm global virulence regulation system. Activated GacA increases the expression of the functional RNA *rsmB,* which down regulates RsmA, allowing translation of RsmA targeted mRNAs.

The Gac system is the central virulence regulator in P.s.

The GacA/S TCST system is directly or indirectly involved in the regulation of every known virulence factor in *P.s.* including coronatine phytotoxin, EPS, and

the T3SS and its effectors [65]. The GacA/S system achieves this breadth by positively regulating the quorum sensing system, crucial to motility and EPS regulation, and the sigma factors RpoS, RpoN and HrpL (Figure 3b) [65]. The Gac system is activated in the host apoplast by a combination of signals, including low pH, low osmolarity, sucrose or fructose sugars and a lack of complex carbon and nitrogen sources [65, 66]. However the exact signal(s) GacS receives and the mechanism of GacA-GacS interaction remain unknown. Additionally, the molecular means by which GacA/S regulates downstream regulators is largely unknown. Answers to these questions should reveal how a pathogen like *P.s.* relies so heavily on a TCST system that functions in one of two states, inactivated or activated, while coordinating quantitative, analogue outputs like the expression of virulence factors.

The RpfC/G TCST system in *X.c.c.* links cell-to-cell signaling to diverse behavioral changes

Unlike in *P.s.* and *P.c.,* a well-characterized TCST system constitutes the novel cell-to-cell communication system of *X.c.c*. The unique RpfC/RpfG TCST system, described above, is not only central to cell-to-cell communication but also to the regulation of diverse behaviors mediated via the degradation of cyclic di-GMP (Figure 3a). RpfC is a hybrid membrane-bound, bi-functional sensor kinase composed of three distinct functional domains that autophosphorylate and phosphorelay to both activate RpfG and repress DSF biosynthesis. RpfG is a novel cyclic di-GMP phosphodiesterase that degrades the bacterial second

messenger cyclic di-GMP [67, 68]. As mentioned above, cyclic di-GMP has global regulatory functions, but the proteins with which it interacts downstream of RpfG are only beginning to be identified. Thus, the RpfC/G TCST system appears to connect large-scale virulence regulation with cell-to-cell communication. However, the many gaps that exist in the network could, when discovered, change our perception of the importance of RpfC and RpfG in global virulence regulation.

III. T3SS gene expression is regulated by AraC-Type regulators or alternative sigma factors.

AraC-Type Regulators regulate T3SS expression in R.s. and X.c.c.

AraC-type regulators commonly have important virulence regulation roles in association with T3SSs in animal and plant bacterial pathogens. Particularly in enterobacterial pathogens such as *Salmonella, Shigella, Yersinia* and *E. coli,* AraC-like regulators are key activators of T3SS expression. *Pseudomonas aeruginosa* also relies on an AraC-like transcription activator, ExsA, to regulate the T3SS [69]. Interestingly, neither *P.s.* nor *P.c.* use AraC-like proteins to regulate their T3SS, while *R.s.* and *X.c.c.* do. In *R.s.,* the response regulator HrpG responds to contact with the plant cell wall via PrhA and activates the expression of the AraC-like HrpB [28, 70] (Figure 2a). HrpB activates the *hrp* regulon (T3SS) promoters by binding a specific PIP box sequence [71, 72]. HrpB

in *R.s.* also regulates chemotaxis and siderophore expression, and generally coordinates expression of virulence genes, but by unknown means [70][73]. In *Xanthomonas campestris*, the response regulator HrpG also activates the expression of the AraC-like regulator HrpX [61, 71] (Figure 3a). HrpX and HrpB are related, and mutation of HrpX can be complemented by HrpB. While the significance of AraC-like regulators is clear in T3SS regulation, the means by which they influence global virulence regulation has not been fully established.

Alternative Sigma Factors regulate T3SS and more

A major theme across T3SS regulation in phytopathogens is the use of alternative sigma factors. Sigma factors are essential transcription initiation factors that direct RNA polymerase to bind specific promoter regions. In *P.s.* and *P.c.*, the alternative sigma factor responsible for activating the expression of the whole T3SS and the effectors is HrpL [15]. HrpL is a member of the extracytoplasmic function (ECF) sigma factor family, a family significant for controlling transcription during stress responses and morphological changes [74]. In both *P.c.* and *P.s.*, another sigma factor, RpoN/sigma-54 activates the expression of HrpL (Figures 2b and 3b). Additionally, specific NtrC-like transcription activators regulate HrpL expression in both cases: HrpR and HrpS in *P.s.* [44] and HrpS in *Erwinia and Pectobacterium* species [21, 75] *(*reviewed in [76]). In *P.s.* HrpR/S activation of the T3SS might be reinforced by the transcription activator SalA [65], but repressed by Lon protease via protein

degredation by HrpR [77].The alternative sigma factor, PrhI, in *R.s.* is regulated by the signal transducer PrhR/A, activated by contact with the plant cell wall. PrhI then activates expression of the transcription activator PrhJ, which is then responsible for activating the expression of HrpG [78] (Figure 2a). While HrpL and PrhI have been studied in relation to T3SS regulation, whole genome expression profiling studies have begun to expand our knowledge of their regulons beyond T3SS gene expression. Most notably, the regulon of HrpL in *P.s.* pv. *tomato* has recently been expanded to include protein synthesis, metabollic genes [79] and might include the coronatine phytotoxin regulator, CorR [76].

CONCLUDING REMARKS

Phytopathogens coordinate transitions in life-histories and infection strategies by collecting information from the host plant, the environment, and their own population density. These inputs result in a collection of virulence outputs in each species determined by integrated, global regulation networks. The current evidence indicates that virulence regulation networks in the phytopathogens we profile here center on global regulators and cell-density dependent regulation. Complementary TCST systems provide environmental signal inputs to virulence regulation networks. AraC-type regulators and alternative sigma factors, historically significant for T3SS regulation, have global regulatory impacts which

are supplemental to global regulators. Many questions still remain in each species, such as the specific environmental factors that trigger regulators; the use of DNA microarrays and genome sequences is advancing our knowledge quickly. Despite common mechanisms and themes among the phytopathogens we discuss, phylogenetic comparisons of their DNA sequences reveals that each species is more closely related to non-plant pathogen species than they are to each other [80]. Identifying how common mechanisms and horizontally acquired virulence determinants have become incorporated into global regulation networks might be the next step in understanding a phytopathogens' success in infection processes and their own biological niches.

The main question we wish to address in this dissertation is how *P. carotovorum* coordinately deploys stealth virulence determinants, like the T3SS, and brute-force, the T2SS, and what environmental signals are inputs to the virulence network described above. This review allows us to identify key regulators in *P. carotovorum*, such as HrpL, *rsmB*, KdgR, and their downstream outputs, which we will use as the basis for dissecting virulence responses. In the following chapters, we will discuss data that furthers our understanding of how *P. carotovorum* differentially regulates virulence determinants, like the T3SS, during different infection routes using its global virulence network and metabolism signals that are inputs to this network.

LITERATURE CITED:

1. Drancourt, M. and D. Raoult, *Molecular insights into the history of plague.*Microbes Infect, 2002. **4**(1): p. 105-9.

2. Nelson, E.J., et al., *Cholera transmission: the host, pathogen and bacteriophage dynamic.* Nat Rev Microbiol, 2009. **7**(10): p. 693-702.

3. Charkowski, A.O., *Decaying signals: will understanding bacterial-plant communications lead to control of soft rot?* Curr. Opin. Biotechnol., 2009. **20**(2): p. 178-84.

4. Yap, M.N., J.D. Barak, and A.O. Charkowski, *Genomic diversity of Erwinia carotovora subsp. carotovora and its correlation with virulence.* Appl Environ Microbiol, 2004. **70**(5): p. 3013-23.

5. Allen, C., S. Reverchon, and J. Robert-Baudouy, *Nucleotide sequence of the Erwinia chrysanthemi gene encoding 2-keto-3-deoxygluconate permease.* Gene, 1989. **83**(2): p. 233-41.

6. Lagaert, S., T. Belien, and G. Volckaert, *Plant cell walls: Protecting the barrier from degradation by microbial enzymes.* Semin. Cell Dev. Biol., 2009. **20**(9): p. 1064-73.

7. Matsumoto, H., et al., *Comparative study of regulatory mechanisms for pectinase production by Erwinia carotovora subsp. carotovora and Erwinia chrysanthemi.* Mol. Plant Microbe Interact., 2003. **16**(3): p. 226-37.

8. Rojas, C.M., et al., *HecA, a member of a class of adhesins produced by diverse pathogenic bacteria, contributes to the attachment, aggregation, epidermal cell killing, and virulence phenotypes of Erwinia chrysanthemi EC16 on Nicotiana clevelandii seedlings.* Proc Natl Acad Sci U S A, 2002. **99**(20): p. 13142-7.

9. Marits, R., et al., *Regulation of the expression of prtW::gusA fusions in Erwinia carotovora subsp. carotovora.* Microbiology, 2002. **148**(Pt 3): p. 835-42.

10. Rantakari, A., et al., *Type III secretion contributes to the pathogenesis of the soft-rot pathogen Erwinia carotovora: partial characterization of the hrp gene cluster.* Mol Plant Microbe Interact, 2001. **14**(8): p. 962-8.

11. Gophna, U., E.Z. Ron, and D. Graur, *Bacterial type III secretion systems are ancient and evolved by multiple horizontal-transfer events.* Gene, 2003. **312**: p. 151-63.

12. Nomura, K., M. Melotto, and S.Y. He, *Suppression of host defense in compatible plant-Pseudomonas syringae interactions.* Curr Opin Plant Biol, 2005. **8**(4): p. 361-8.

13. Belkhadir, Y., R. Subramaniam, and J.L. Dangl, *Plant disease resistance protein signaling: NBS-LRR proteins and their partners.* Curr Opin Plant Biol, 2004. **7**(4): p. 391-9.

14. Chang, J.H., et al., *A high-throughput, near-saturating screen for type III effector genes from Pseudomonas syringae.* Proc. Natl. Acad. Sci. U.S.A., 2005. **102**(7): p. 2549-54.

15. Chatterjee, A., Y. Cui, and A.K. Chatterjee, *Regulation of Erwinia carotovora hrpL(Ecc) (sigma-L(Ecc)), which encodes an extracytoplasmic function subfamily of sigma factor required for expression of the HRP regulon.* Mol Plant Microbe Interact, 2002. **15**(9): p. 971-80.

16. Toth, I.K. and P.R. Birch, *Rotting softly and stealthily.* Curr Opin Plant Biol, 2005. **8**(4): p. 424-9.

17. Barnard, A.M. and G.P. Salmond, *Quorum sensing in Erwinia species.* Anal Bioanal Chem, 2007. **387**(2): p. 415-23.

18. Cui, Y., et al., *rsmC of the soft-rotting bacterium Erwinia carotovora subsp. carotovora negatively controls extracellular enzyme and harpin(Ecc) production and virulence by modulating levels of regulatory RNA (rsmB) and RNA-binding protein (RsmA).* J Bacteriol, 1999. **181**(19): p. 6042-52.

19. Liu, Y., et al., *Characterization of a novel RNA regulator of Erwinia carotovora ssp. carotovora that controls production of extracellular enzymes and secondary metabolites.* Mol Microbiol, 1998. **29**(1): p. 219-34.

20. Mukherjee, A., et al., *hexA of Erwinia carotovora ssp. carotovora strain Ecc71 negatively regulates production of RpoS and rsmB RNA, a global regulator of extracellular proteins, plant virulence and the quorum-sensing signal, N-(3-oxohexanoyl)-L-homoserine lactone.* Environ Microbiol, 2000. **2**(2): p. 203-15.

21. Wei, Z., J.F. Kim, and S.V. Beer, *Regulation of hrp genes and type III protein secretion in Erwinia amylovora by HrpX/HrpY, a novel two-*

component system, and HrpS. Mol Plant Microbe Interact, 2000. **13**(11): p. 1251-62.

22. Schell, M.A., *Control of Virulence and Pathogenicity Genes of Ralstonia Solanacearum by an Elaborate Sensory Network.* Annu Rev Phytopathol, 2000. **38**: p. 263-292.

23. Crossman, L. and J.M. Dow, *Biofilm formation and dispersal in Xanthomonas campestris.* Microbes Infect, 2004. **6**(6): p. 623-9.

24. Brumbley, S.M., B.F. Carney, and T.P. Denny, *Phenotype conversion in Pseudomonas solanacearum due to spontaneous inactivation of PhcA, a putative LysR transcriptional regulator.* J Bacteriol, 1993. **175**(17): p. 5477-87.

25. Clough, S.J., et al., *A two-component system in Ralstonia (Pseudomonas) solanacearum modulates production of PhcA-regulated virulence factors in response to 3-hydroxypalmitic acid methyl ester.* J Bacteriol, 1997. **179**(11): p. 3639-48.

26. Flavier, A.B., et al., *Identification of 3-hydroxypalmitic acid methyl ester as a novel autoregulator controlling virulence in Ralstonia solanacearum.* Mol Microbiol, 1997. **26**(2): p. 251-9.

27. Huang, J., et al., *Joint transcriptional control of xpsR, the unusual signal integrator of the Ralstonia solanacearum virulence gene regulatory network, by a response regulator and a LysR-type transcriptional activator.* J Bacteriol, 1998. **180**(10): p. 2736-43.

28. Genin, S., et al., *Control of the Ralstonia solanacearum Type III secretion system (Hrp) genes by the global virulence regulator PhcA.* FEBS Lett, 2005. **579**(10): p. 2077-81.

29. Bhatt, G. and T.P. Denny, *Ralstonia solanacearum iron scavenging by the siderophore staphyloferrin B is controlled by PhcA, the global virulence regulator.* J Bacteriol, 2004. **186**(23): p. 7896-904.

30. Flavier, A.B., et al., *Hierarchical autoinduction in Ralstonia solanacearum: control of acyl-homoserine lactone production by a novel autoregulatory system responsive to 3-hydroxypalmitic acid methyl ester.* J Bacteriol, 1997. **179**(22): p. 7089-97.

31. Pirhonen, M., et al., *A small diffusible signal molecule is responsible for the global control of virulence and exoenzyme production in the plant pathogen Erwinia carotovora.* Embo J, 1993. **12**(6): p. 2467-76.

32. Welch, M., et al., *N-acyl homoserine lactone binding to the CarR receptor determines quorum-sensing specificity in Erwinia.* Embo J, 2000. **19**(4): p. 631-41.

33. Sjoblom, S., et al., *Cooperation of two distinct ExpR regulators controls quorum sensing specificity and virulence in the plant pathogen Erwinia carotovora.* Mol Microbiol, 2006. **60**(6): p. 1474-89.

34. Cui, Y., et al., *ExpR, a LuxR homolog of Erwinia carotovora subsp. carotovora, activates transcription of rsmA, which specifies a global regulatory RNA-binding protein.* J Bacteriol, 2005. **187**(14): p. 4792-803.

35. White, D., M.E. Hart, and T. Romeo, *Phylogenetic distribution of the global regulatory gene csrA among eubacteria.* Gene, 1996. **182**(1-2): p. 221-3.

36. Whitehead, N.A., et al., *The regulation of virulence in phytopathogenic Erwinia species: quorum sensing, antibiotics and ecological considerations.* Antonie Van Leeuwenhoek, 2002. **81**(1-4): p. 223-31.

37. Baker, C.S., et al., *CsrA regulates glycogen biosynthesis by preventing translation of glgC in Escherichia coli.* Mol Microbiol, 2002. **44**(6): p. 1599-610.

38. Mukherjee, A., et al., *RpoS (sigma-S) controls expression of rsmA, a global regulator of secondary metabolites, harpin, and extracellular proteins in Erwinia carotovora.* J Bacteriol, 1998. **180**(14): p. 3629-34.

39. Cui, Y., et al., *Identification of a global repressor gene, rsmA, of Erwinia carotovora subsp. carotovora that controls extracellular enzymes, N-(3-oxohexanoyl)-L-homoserine lactone, and pathogenicity in soft-rotting Erwinia spp.* J Bacteriol, 1995. **177**(17): p. 5108-15.

40. Romeo, T., *Global regulation by the small RNA-binding protein CsrA and the non-coding RNA molecule CsrB.* Mol Microbiol, 1998. **29**(6): p. 1321-30.

41. Cui, Y., A. Chatterjee, and A.K. Chatterjee, *Effects of the two-component system comprising GacA and GacS of Erwinia carotovora subsp. carotovora on the production of global regulatory rsmB RNA, extracellular enzymes, and harpinEcc.* Mol Plant Microbe Interact, 2001. **14**(4): p. 516-26.

42. Liu, Y., et al., *kdgREcc negatively regulates genes for pectinases, cellulase, protease, HarpinEcc, and a global RNA regulator in Erwinia carotovora subsp. carotovora.* J Bacteriol, 1999. **181**(8): p. 2411-21.

43. Valverde, C., et al., *A repeated GGA motif is critical for the activity and stability of the riboregulator RsmY of Pseudomonas fluorescens.* J Biol Chem, 2004. **279**(24): p. 25066-74.

44. Wang, L.H., et al., *A bacterial cell-cell communication signal with cross-kingdom structural analogues.* Mol Microbiol, 2004. **51**(3): p. 903-12.

45. He, Y.W., et al., *Genome scale analysis of diffusible signal factor regulon in Xanthomonas campestris pv. campestris: identification of novel cell-cell communication-dependent genes and functions.* Mol Microbiol, 2006. **59**(2): p. 610-22.

46. Dow, J.M., et al., *Biofilm dispersal in Xanthomonas campestris is controlled by cell-cell signaling and is required for full virulence to plants.* Proc Natl Acad Sci U S A, 2003. **100**(19): p. 10995-1000.

47. Slater, H., et al., *A two-component system involving an HD-GYP domain protein links cell-cell signalling to pathogenicity gene expression in Xanthomonas campestris.* Mol Microbiol, 2000. **38**(5): p. 986-1003.

48. Ross, P., et al., *Regulation of cellulose synthesis in Acetobacter xylinum by cyclic diguanylic acid.* Nature, 1987. **325**: p. 279-281.

49. Romling, U. and D. Amikam, *Cyclic di-GMP as a second messenger.* Curr Opin Microbiol, 2006. **9**(2): p. 218-28.

50. Fouhy, Y., et al., *Cell-cell signaling, cyclic di-GMP turnover and regulation of virulence in Xanthomonas campestris.* Res Microbiol, 2006. **157**(10): p. 899-904.

51. He, Y.W., et al., *Xanthomonas campestris cell-cell communication involves a putative nucleotide receptor protein Clp and a hierarchical signalling network.* Mol Microbiol, 2007.

52. Hsiao, Y.M., et al., *Clp and RpfF up-regulate transcription of pelA1 gene encoding the major pectate lyase in Xanthomonas campestris pv. campestris.* J Agric Food Chem, 2009. **57**(14): p. 6207-15.

53. Quinones, B., C.J. Pujol, and S.E. Lindow, *Regulation of AHL production and its contribution to epiphytic fitness in Pseudomonas syringae.* Mol Plant Microbe Interact, 2004. **17**(5): p. 521-31.

54. Quinones, B., G. Dulla, and S.E. Lindow, *Quorum sensing regulates exopolysaccharide production, motility, and virulence in Pseudomonas syringae.* Mol Plant Microbe Interact, 2005. **18**(7): p. 682-93.

55. Monier, J.M. and S.E. Lindow, *Differential survival of solitary and aggregated bacterial cells promotes aggregate formation on leaf surfaces.* Proc Natl Acad Sci U S A, 2003. **100**(26): p. 15977-82.

56. Beier, D. and R. Gross, *Regulation of bacterial virulence by two-component systems.* Curr Opin Microbiol, 2006. **9**(2): p. 143-52.

57. Huang, J., et al., *A complex network regulates expression of eps and other virulence genes of Pseudomonas solanacearum.* J Bacteriol, 1995. **177**(5): p. 1259-67.

58. Garg, R.P., et al., *Multicomponent transcriptional regulation at the complex promoter of the exopolysaccharide I biosynthetic operon of Ralstonia solanacearum.* J Bacteriol, 2000. **182**(23): p. 6659-66.

59. Allen, C., J. Gay, and L. Simon-Buela, *A regulatory locus, pehSR, controls polygalacturonase production and other virulence functions in Ralstonia solanacearum.* Mol Plant Microbe Interact, 1997. **10**(9): p. 1054-64.

60. Flego, D., et al., *A two-component regulatory system, pehR-pehS, controls endopolygalacturonase production and virulence in the plant pathogen Erwinia carotovora subsp. carotovora.* Mol Plant Microbe Interact, 2000. **13**(4): p. 447-55.

61. Wengelnik, K., G. Van den Ackerveken, and U. Bonas, *HrpG, a key hrp regulatory protein of Xanthomonas campestris pv. vesicatoria is homologous to two-component response regulators.* Mol Plant Microbe Interact, 1996. **9**(8): p. 704-12.

62. Schneider, P., et al., *The global virulence regulators VsrAD and PhcA control secondary metabolism in the plant pathogen Ralstonia solanacearum.* Chembiochem, 2009. **10**(17): p. 2730-2.

63. Laville, J., et al., *Global control in Pseudomonas fluorescens mediating antibiotic synthesis and suppression of black root rot of tobacco.* Proc Natl Acad Sci U S A, 1992. **89**(5): p. 1562-6.

64. Hrabak, E.M. and D.K. Willis, *The lemA gene required for pathogenicity of Pseudomonas syringae pv. syringae on bean is a member of a family of two-component regulators.* J Bacteriol, 1992. **174**(9): p. 3011-20.

65. Chatterjee, A., et al., *GacA, the response regulator of a two-component system, acts as a master regulator in Pseudomonas syringae pv. tomato DC3000 by controlling regulatory RNA, transcriptional activators, and alternate sigma factors.* Mol Plant Microbe Interact, 2003. **16**(12): p. 1106-17.

66. Rahme, L.G., M.N. Mindrinos, and N.J. Panopoulos, *Plant and environmental sensory signals control the expression of hrp genes in Pseudomonas syringae pv. phaseolicola.* J Bacteriol, 1992. **174**(11): p. 3499-507.

67. Dow, J.M., et al., *The HD-GYP domain, cyclic di-GMP signaling, and bacterial virulence to plants.* Mol Plant Microbe Interact, 2006. **19**(12): p. 1378-84.

68. Ryan, R.P., et al., *Cell-cell signaling in Xanthomonas campestris involves an HD-GYP domain protein that functions in cyclic di-GMP turnover.* Proc Natl Acad Sci U S A, 2006. **103**(17): p. 6712-7.

69. Francis, M.S., H. Wolf-Watz, and A. Forsberg, *Regulation of type III secretion systems.* Curr Opin Microbiol, 2002. **5**(2): p. 166-72.

70. Valls, M., S. Genin, and C. Boucher, *Integrated regulation of the type III secretion system and other virulence determinants in Ralstonia solanacearum.* PLoS Pathog, 2006. **2**(8): p. e82.

71. Koebnik, R., et al., *Specific binding of the Xanthomonas campestris pv. vesicatoria AraC-type transcriptional activator HrpX to plant-inducible promoter boxes.* J Bacteriol, 2006. **188**(21): p. 7652-60.

72. Mukaihara, T., et al., *Genetic screening of Hrp type III-related pathogenicity genes controlled by the HrpB transcriptional activator in Ralstonia solanacearum.* Mol Microbiol, 2004. **54**(4): p. 863-75.

73. Occhialini, A., et al., *Genome-wide analysis of gene expression in Ralstonia solanacearum reveals that the hrpB gene acts as a regulatory switch controlling multiple virulence pathways.* Mol Plant Microbe Interact, 2005. **18**(9): p. 938-49.

74. Helmann, J.D., *The extracytoplasmic function (ECF) sigma factors.* Adv Microb Physiol, 2002. **46**: p. 47-110.

75. Hutcheson, S.W., et al., *Enhancer-binding proteins HrpR and HrpS interact to regulate hrp-encoded type III protein secretion in Pseudomonas syringae strains.* J Bacteriol, 2001. **183**(19): p. 5589-98.

76. Tang, X., Y. Xiao, and J.M. Zhou, *Regulation of the type III secretion system in phytopathogenic bacteria.* Mol Plant Microbe Interact, 2006. **19**(11): p. 1159-66.

77. Losada, L.C. and S.W. Hutcheson, *Type III secretion chaperones of Pseudomonas syringae protect effectors from Lon-associated degradation.* Mol Microbiol, 2005. **55**(3): p. 941-53.

78. Brito, B., et al., *A signal transfer system through three compartments transduces the plant cell contact-dependent signal controlling Ralstonia solanacearum hrp genes.* Mol Plant Microbe Interact, 2002. **15**(2): p. 109-19.

79. Ferreira, A.O., et al., *Whole-genome expression profiling defines the HrpL regulon of Pseudomonas syringae pv. tomato DC3000, allows de novo reconstruction of the Hrp cis clement, and identifies novel coregulated genes.* Mol Plant Microbe Interact, 2006. **19**(11): p. 1167-79.

80. de Souza, J.T., M. Mazzola, and J.M. Raaijmakers, *Conservation of the response regulator gene gacA in Pseudomonas species.* Environ Microbiol, 2003. **5**(12): p. 1328-40.

CHAPTER 2 : Niche-specificity and the variable fraction of the *Pectobacterium* pan-genome

ABSTRACT:

In this chapter, we compare genome sequences of three closely related soft rot pathogens that vary in host range and geographical distribution to identify genetic differences that could account for lifestyle differences. This work was done in collaboration with Jeremy Glasner, Nicole Perna, and Amy Charkowski at the University of Wisconsin. My contribution to this work includes preparing DNA that was used for genome sequencing, assisting with the annotation of the genome, and manuscript preparation. The isolates compared, *Pectobacterium atrosepticum* SCRI1043, *Pectobacterium carotovorum* WPP14 and *Pectobacterium brasiliensis* 1692, represent diverse lineages of the genus. *Pc* and *Pb* genome contigs generated by 454 pyrosequencing ordered by reference to the previously published complete circular chromosome of *Pa* genome, and each other, account for 96% of the predicted genome size. Orthologous proteins encoded by *Pc* and *Pb* are about 95% identical to each other and 92% identical to *Pa*. Multiple alignment using Mauve identified a core genome of 3.9 Mb conserved among these *Pectobacterium* species. Each core genome is interrupted at many points by species-specific insertions or deletions (indels) that account for about 0.9-1.1 Mb. We demonstrate that the presence of a *hrpK*-like type III secretion system dependent effector protein in *Pc* and *Pb,* and its absence from *Pa* is insufficient to explain variability in their response to infection in a plant. Additional genes that vary among these species include those encoding peptide toxin production, enzyme production, secretion proteins, and

antibiotic production, as well as differences in more general aspects of gene regulation and metabolism that may be relevant to pathogenicity.

INTRODUCTION

Comparative genomics can reveal physiological and functional variation among bacteria that provides insight into their ability to exploit distinct ecological niches. For plant pathogenic bacteria, key factors related to interaction with a plant host have emerged from comparisons with closely related animal-associated bacteria. For example, identification of gene clusters, or islands, that distinguish a single plant-pathogenic *Pectobacterium* genome from those of other enterobacteria that infect animals, like *Escherichia* and *Salmonella*, revealed known and novel virulence factors, many of which are also found in more distantly related plant pathogens [1, 2]. Comparison of more closely related genomes can be particularly useful for illuminating niche adaptation, in part simply because the level of observed genetic variation is lower, simplifying reconstruction of both the evolutionary history and the phenotypic consequences of individual polymorphisms. Intra-specific comparisons of genomes from strains of both *Pseudomonas syringae* and *Xanthomonas campestris* with diverse host ranges revealed differences including the type and number of type III secreted proteins [3-5]. These can affect host range either by suppressing host defenses, enabling the growth of a pathogen, or conversely, they can trigger a strong resistance

response blocking pathogen growth if an individual type III secreted protein is recognized by a cognate resistance protein in the host [6].

There are no previous intra-specific genome sequence comparisons for plant-associated enterobacteria, but several groups of animal-associated enterobacteria have been sampled fairly extensively. Early comparisons revealed that only 40% of the total distinct protein coding sequences are shared among the model *Escherichia coli* K12 strain, an enterohemorrhagic strain, and a uropathogenic strain [7] despite the fact that over 3.5 Mb of the 4.5-5.5 Mb core genome is conserved among all three. This pattern has been observed in other groups of bacteria [8, 9] leading to the concept of a pan-genome, or the complete collection of genes in the species, only a fraction of which are found in any given strain. Importantly, as additional genomes (lineages) of a species are added to comparisons, the fraction of genes conserved, or core genome, declines relatively slowly for most bacteria, reflecting the relative stability of core metabolic and information processing related gene-content. In contrast, some "cosmopolitan" species, like *E. coli* and *Streptococcus agalactiae* have a much larger and dynamic pan-genome, with the variable fraction composed of the sum total of many distinct lineage-specific gene clusters, or islands, acquired through horizontal gene transfer. Mathematical models predict that even relatively well-sampled lineages like *E. coli* will continue to yield novel genes as more strains are sequenced [10]. For plant-pathogenic enterobacteria, with one published genome sequence, characterization of the pan-genome is in its infancy.

The soft rot pectobacteria (previously known as *Erwinia carotovora*) are economically important plant pathogenic enterobacteria. *Pectobacterium* species cause a spectrum of disease symptoms, termed wilt, soft rot, and blackleg on a wide range of monocot and dicot host plants. These diseases are responsible for large economic losses during potato and ornamental production. *Pectobacterium* have been isolated from numerous plant hosts, soil, and both surface and ground water [11, 12]. *Pectobacterium* have also been found in association with a variety of invertebrates, ranging from fruit flies to snails [13-15]. These invertebrates play an important role in the spread of soft rot bacteria between plant hosts and may also respond to the plant pathogen. For example, Basset et al. [16] showed that fruit flies have an immune response when inoculated with *Pectobacterium*.

The genus *Pectobacterium* was recently resurrected and four *Erwinia carotovora* subspecies were renamed as different *Pectobacterium* species; *P. atrosepticum*, *P. betavasculorum*, *P. carotovorum*, and *P. wasabiae* [17]. Concurrently, a potential fifth species, originally named *E. carotovora* subsp. *brasiliensis* was described as causing blackleg on potatoes in Brazil [18]. This species is phylogenetically distinct from the other four *Pectobacterium* species [19] and will be referred to herein as *P. brasiliensis.* The pathogens *P. carotovorum* [20], *P. atrosepticum* [21], *P. brasiliensis* [22] and *P. wasabiae* [23] all cause disease on potato. Multiple species may cause disease in the same field, and even on the same individual plant. *Pc* is found in many climates worldwide, *Pa* is found in

cool climates worldwide, and *Pb* has only been reported in Brazil, Israel, and the United States, but is likely to have a wider distribution [19]. Of the potato-infecting species, *Pc*, *Pb*, and *Pw* have all been reported to cause disease on other plant species and are thus broad host range pathogens, while *Pa* appears to be limited to potato and closely related solanaceous crops. The genetic differences that limit the host range of *Pa* in comparison to other *Pectobacterium* species are unknown.

In order to identify variation between subspecies that could provide clues to the nature of differences in host range, nutritional requirements and species niches in the *Pectobacteria,* we compared the genome sequences of three isolates that were phlyogentically well distributed. A phylogenetic analysis based on sequencing fragments of seven housekeeping genes from bacterial soft rot isolates collected from diverse locations around the world revealed that the *Pectobacterium* species could be classified into a monophyletic group distinct from the broad host range soft rot pathogen *Dickeya* [24] and the tree pathogen *Brenneria*. The pectobacteria could be divided into 5 clades related by monophyletic descent [19]. We chose the previously sequenced strain *Pa* SCRI1043 representing Clade V, *Pc* WPP14 representing Clade II and *Pb* 1692 representing Clade I. Conserved genes from these species were, on average, 95% identical. This degree of homology allowed us to assemble multiple sequence contigs derived by 454 platform pyrosequenceing (Roche life sciences) and Newbler analysis of genomic DNA from *Pb* and *Pc* into ordered genome

sequences by alignment with the genome of *Pa* SCRI1043. Although we used the *Pa* SCRI1043 genome to establish the structure of the other two genomes, the depth of sequence economically available with pyrosequencing allowed us to identify and position genes unique to each species as well as conserved sequences.

Comparison of the three genomes revealed a common core genome, (representing approximately 80% of the nucleotides in each species) dotted with islands carrying diverse sequences. The genes in unique islands were enriched for proteins of DNA replication, mostly of phage origin, and in regulatory genes. Unlike the xanthomonads and *Pseudomonas syringae*, and even the blight causing *Erwinia amylovora,* [25] the predicted type III secreted proteins were highly conserved within the group. Variability that could be associated with differences in host range was seen in genes for type IV secretion systems, putative phytotoxins, taxis and motility genes, and cell surface proteins.

MATERIALS AND METHODS

Genomic analysis of *Pc* and *Pb*

Pc strain WPP14 [26] and *Pb* strain 1692 [18] were streaked from frozen cultures onto LB plates. Genomic DNA was isolated [27] from cultures of *Pc* and *Pb* grown in LB broth in shaking incubators at 25 and 37 degrees Celsius respectively. *Pc* was sequenced in three standard runs on a 454 GS20 by 454

Life Sciences (Branford, CT) and assembled with Newbler software version 1.0.51.02. *Pb* was sequenced in two standard runs on a GS20 by the Washington University Genome Center and assembled with Newbler software version 1.0.53.12. Contigs were reordered using Projector 2.0 [28] and Mauve 2.0 [29] which was also employed for multiple genome alignment. Primer pairs for 188 gaps in the *Pc* genome sequence were designed using Projector 2.0 and used in PCR with *Pc* genomic DNA. Selected PCR products were sequenced by standard dye-deoxy sequencing on an ABI capillary sequencer.

We used a combination of TBLASTN [30] searches with predicted proteins from *Pa* SCRI1043 [1] and Glimmer2 [31] to predict ORFs in the *Pc* and *Pb* genomes. Manual inspection of predicted ORF was used to remove some ORFs that overlapped other ORFs and small ORFs (<250 bp) that lacked significant BLASTP (E-value <0.00001) matches against the GenPept database. We used InterProScan [23] to identify protein domains in the *Pectobacterium* genomes and obtained Gene Ontology (GO) (Ashburner et al. 2000) terms associated with each domain from the InterPro to GO term mapping available from the GO website.

Sequence data and analyses are available for download from the ASAP database [32].

Expression of *Pc hrpK* in *Pa*

We PCR amplified the *hrpK* gene from *Pc*, including a 135 bp upstream region relative to the translation start site of *hrpK* that contains a putative HrpL-binding motif, using primers P0492 (5'-taa gag tca gga gct agt gtg gcg gag ctc agg gtt-3') and P0493 (5'-taa gct ggc gca tta gcg cga att cgg aat att g-3'). The PCR frogment was cloned into pCPP50 [33], resulting in plasmid p50*hrpK*. This plasmid was electroporated into *Pa* and a single ampicillin resistant colony was isolated for hypersensitive response assays.

Deletion of *hrpK* from *Pc*

Regions flanking the *Pc hrpK* gene were PCR amplified with the left primer set P0504 (5'- gtg ctg gat ccg cta ata tca tca tac-3') and P0505 (5'-cgt act ctg cga agc ttc ccg tcc cca ttc tgc tgt tgt ca-3') and right primer set P0506 (5'-gga agc ttc gca gag tac gat tcc caa acc gcg cta atg c-3') and P0507 (5'-gtc tgc cgg atc cac gtt taa cga t-3'). These two PCR fragments were used for as templates for crossover PCR with primers P0504 and P0507. The 2.9-kb product was cloned into pGEMT-easy (Promega, Madison, WI), resulting in a plasmid pTA∆*hrpK*_ABCD. A chloramphenicol resistance and GFP cassette obtained from pTAgfp::cm was cloned into the *Hin*dIII site of pTA∆*hrpK*_ABCD, producing pTA∆*hrpK*::gfp::cm This plasmid was electroporated into *Pc* for allelic-exchange mutagenesis following the methods described by Ried and Collmer [34]. The *hrpK* deletion mutation was confirmed by PCR and Southern blot analysis.

Methods for electroporation, restriction endonuclease digestion, PCR, cloning and Southern blotting were performed as described in Sambrook et al. [27].

Hypersensitive response assay in *Nicotiana tabacum*

An assay for the hypersensitive response [17] was performed essentially as described in Bauer et al. (Bauer et al. 1994). Six to seven week old *N. tabacum* cv Xanthi leaves were infiltrated with either *Pa*, *Pa* with p50*hrpK*, *Pc, Pc*Δ*hrpK* or water as a negative control. Plants were visually assayed for HR elicitation after 24 hours.

RESULTS AND DISCUSSION

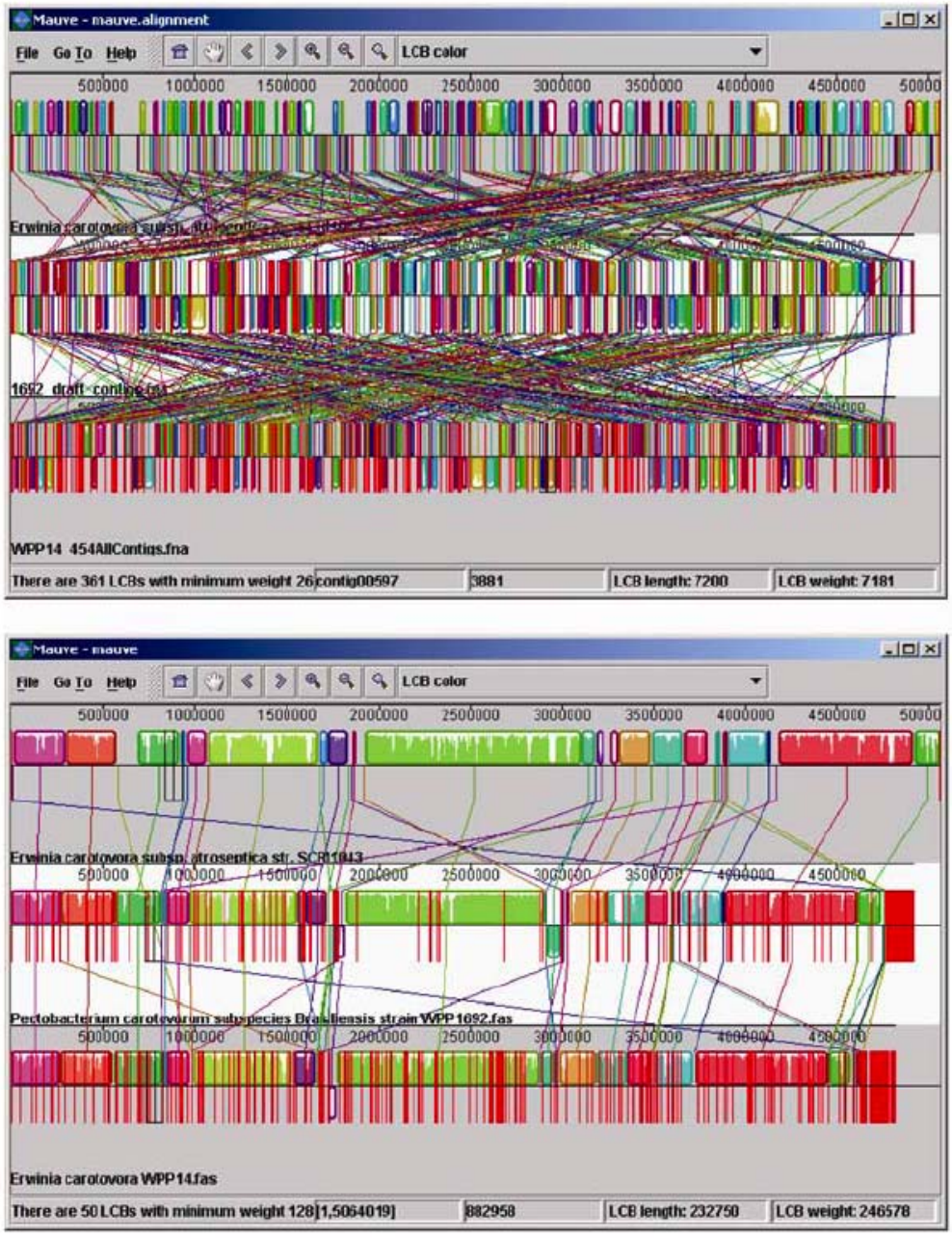
We used 454-platform massively parallel pyrosequencing [35] to generate draft genome sequences for *Pc* and *Pb*. Comparisons to a published complete genome for *Pa*, and to each other allowed us to order and orient the large number of contigs onto a scaffold based on the *Pa* sequence, thereby constructing virtual genomes for *Pc* and *Pb* and facilitating examination of higher order features, such as the presence, absence, or rearrangement of gene islands (Figure 1).

Assembly and ordering contigs in the draft genomes. Summary statistics are shown in Table 1. Three runs of the 454 GS20 instrument on a DNA preparation

for *Pc* resulted in 116,320,270 bp of usable sequence, or 23.3-fold coverage of the genome assuming a final size of 5 Mb, similar to *Pa* and most other enterobacteria. Two runs using the same instrumentation yielded 86,099,090 bp and level of 17.2-fold coverage for *Pb*. *De novo* assembly for each genome was performed using the standard Newbler assembler resulting in 4,746,006 bp and 4,775,163 bp in contigs greater than 500 bp in length for *Pc* and *Pb*, respectively. We employed two approaches to predict the order and orientation of the large number of contigs from each draft genome, Projector 2.0 [28] and Mauve 2.0 [29].

Figure 1. Mauve alignments of the *Pa* genome and the draft genomes of *Pb* and *Pc* before (top image) and after (bottom image) contig reordering. Each alignment has three panels, one for each genome (*Pa*, *Pb*, *Pc*), composed of colored segments corresponding to the boundaries of locally collinear blocks with lines connecting the center of homologous blocks in each genome. Vertical red lines in the *Pb* and *Pc* panels indicate contig boundaries. A significant number of small, unordered lineage-specific contigs appear as a dense red region at the end of the *Pb* and *Pc* genome panels. The reduction in the number of locally collinear blocks from 361 in the top alignment to 50 in the bottom alignment and increase in average block length is clearly visible.

Figure 1.



Projector 2.0 uses BLAST searches of strategic segments of each contig against a complete (or draft) reference genome to predict their relative locations and also outputs a list of PCR primer pairs designed to link adjacent contigs. We tested

188 pairs of primers with *Pc* and confirmed 84 predicted contig linkages (data not shown). It is likely that more could be confirmed by optimizing PCR conditions.

Upon aligning the Projector 2.0 ordered contig sets for both draft genomes with the complete *Pa* genome, we observed that Mauve detected additional contigs that could be ordered, including several cases where the *Pc* and *Pb* contigs could be ordered relative to each other even if they were not homologous to regions of *Pa*. Starting with the Projector 2.0 ordered set, we iteratively aligned and reordered contigs until no additional members of existing locally collinear blocks were found. After running both Projector 2.0 and Mauve, order and orientation can be predicted for a total of 273/731 *Pc* contigs and 169/1370 *Pb* contigs. These ordered contigs account for 96.6% and 96.8% of their respective genomes. Figure 1 compares Mauve alignments for the unordered and ordered contig sets. The final alignment of ordered contigs is available through ASAP and was used for a variety of the analyses described below.

Comparison of *Pc* draft genome with physical chromosome map. A preliminary physical map of the *Pc* chromosome was constructed using restriction enzyme mapping, mutagenesis, and DNA hybridization [26] and this map was compared to the assembled draft sequence. Yap et al. [26] digested *Pc* genomic DNA with I-*Ceu*I, which recognizes a conserved sequence in bacterial 23S rRNA genes, and observed seven I-*Ceu*I fragments by pulsed field gel

Table 1. Summary statistics for three sequenced *Pectobacterium* genomes[a]

Genome	*P. atrosepticum* SCRI1043	*P. brasiliensis* 1692	*P. carotovorum* WPP14
Sequencing method	Dye terminator chemistry on ABI3700 automated sequencers	Two plates of 454 GS20 sequencing	Three plates of 454 GS20 sequencing
Assembly method	assembled using PHRAP and GAP4 for finishing	Newbler (version 1.0.53.12)	Newbler (version 01.0.51.02)
Total number of reactions	54,600 paired-end reads (two small insert libraries); 25,800 paired-end reads (medium insert library); 500 paired-end reads (large insert library)	804,834	1,080,725
Number of reads assembled	80,900	782,719	1,057,457
Approximate read length	500	110	110
Total number of bases sequenced	40,450,000	86,099,090	116,320,270
Finished sequence	Final closed sequence composed of 106,500 reads; average 10.2-fold coverage	Unclosed sequence with approximately 17.2-fold coverage	Unclosed sequence with approximately 23.3-fold coverage
Total number of contigs	1	1,370	731
Total base pairs	5,064,019	4,918,574	4,823,187
Number of large contigs (≈>500 bp)	1	143	272
Total bp in large contigs	5,064,019	4,775,163	4,746,006
Average contig size	na	34,353	17,775
N50 contig size	na	75,304	38,023
Largest contig size	na	350,808	102,016
Number of predicted proteins	4,492	4,387	4,245
Number of pseudogenes identified	21	nd	150
Reported	Bell et al. 2004	This study	This study

[a] Abbreviations: na = not available and nd = not determined.

electrophoresis (PFGE), suggesting that *Pc*, like most *Pectobacterium* strains, encodes seven rRNA operons. Sizes of the smallest six fragments were estimated and compared to the draft genome alignment. All six fragments could be identified in the draft genome and all were roughly the same size as observed by PFGE, thus the draft genome sequence appears to represent the majority of the genome sequence. The physical map obtained by Yap et al. [26] also correctly represented the order of the seven I-*Ceu*I fragments in the WPP14 genome.

Figure 2.

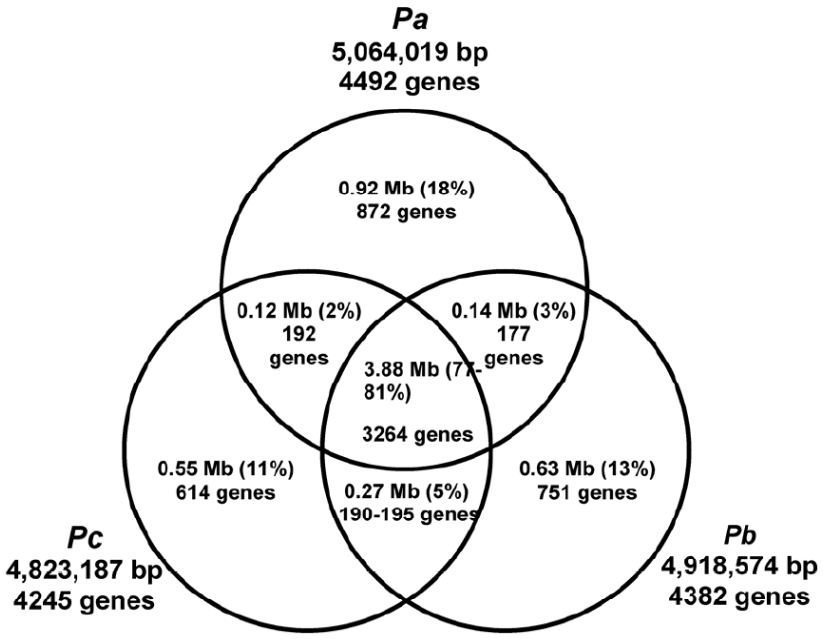


Figure 2. Venn diagram illustrating the total amounts of shared and lineage-specific nucleotide sequence observed in the Mauve multiple alignment and predicted orthologous genes. Gene fragments resulting from likely sequence errors in the two draft genomes leads to gene count differences between species even within orthologous regions. For simplicity, the gene count for *Pa* is shown for all shared regions except the comparison between *Pc* and *Pb*, where counts are shown for both species.

Gene prediction, annotation, and comparative genomics. In order to minimize redundant effort and maximize consistency across genomes, we used a comparative approach to genome annotation. TBLASTN searches using the predicted proteins from the complete *Pa* genome as queries against the draft genomes, were used to identify and annotate the boundaries of all intact open

reading frames. This analysis also revealed many gene fragments that would be annotated as pseudogenes if they occurred in complete genomes; however, we expect that in these draft genomes, most of these disrupted reading frames are caused by sequencing errors arising from both known [35] and potentially uncharacterized sources of error associated with pyrosequencing and/or assembly. PCR and resequencing using traditional Sanger sequencing chemistry confirmed that 10 out of 10 examples were sequencing errors (data not shown). We filtered Glimmer 2 results to remove predicted open reading frames that overlapped the set inherited from *Pa* in order to identify lineage-specific ORFs. Orthologs were predicted using pair-wise reciprocal BLASTP searches retaining only best hits that include greater than 60% of both predicted proteins and show greater than 55% identity. Proteins with a second match that had an *E-value* within one order of magnitude were labeled homologs rather than orthologs. Additionally, shorter (30-60% aligned) matches with comparable levels of sequence identity to orthologs were labeled homologs. This allowed us to accommodate gene fragments. Annotations for *Pc* and *Pb* genes with *Pc* orthologs were propagated directly without further manual review. Annotations for lineage-specific genes were generated using the RAST server (http://rast.nmpdr.org/). Genes found exclusively in *Pc* or *Pb* and genes found in both, but absent from *Pa*, that also had a significant match (*E-value* > 0.00001) to a potentially informative (did not contain “unknown”, “hypothetical”, “phage”, “orf”, ”transposase”) GenPept entry, were flagged for manual review by human

experts. All annotations, automated and manual, and orthologs relationships, were managed and distributed using ASAP [36].

Alignment of the three *Pectobacterium* genomes using Mauve (Figures 1 and 2) revealed that approximately 77% of the complete *Pa* chromosome is present in *Pb* and *Pc*. This is comparable to the fraction of core genome relative to the total pan-genome in the previous three *E. coli* strain comparison [7], but the variable fraction of the pan-genome is more evenly distributed among the three *Pectobacterium* strains and subset genome pairs. Roughly 5.4 % of the *Pb* and *Pc* sequences match each other, but are not conserved with *Pa*. This is close to double the amount of sequence shared by *Pa* and either *Pb* or *Pc*, but not both, supporting the closer relationship of *Pb* and *Pc* observed by Ma et al. [19]. This estimate of the size of the core genome based on genome alignment is more reliable than what can be achieved by counting BLASTP-predicted homologs, in part because it includes intergenic regions and in part, because it is more robust with regard to pseudogenes and sequencing errors. An extreme example is illustrated in Fig. S1, which shows a region encoding two large (>7000 amino acids each) predicted components of a non-ribosomal synthase in *Pa*. No homologous proteins are detected in either *Pb* or *Pc* even though DNA sequences homologous to almost the entire 42 kb region are present in both draft genomes. The large number of relatively small contigs found in this region of the draft genomes suggests that this type of highly repetitive gene poses assembly problems with short pyrosequencing reads. Keeping this caveat in

mind, at least 3251 of 4492 predicted *Pa* genes have orthologs (or close homologs) in all three species, and an additional 177 and 192 genes have matches in *Pb* or *Pc*, respectively. Likely sequencing errors are expected to have an even greater impact on inference of orthologs between the two draft genomes, where we detect 190-195 conserved genes. Average amino acid identity between *Pc* and *Pb* is higher (95%) than observed between orthologs of either *Pc* or *Pb*, and *Pa* (92%) underscoring the closer relationship between these two lineages. We strongly suggest that users consult the multiple alignment if they are interested in conservation of a particular gene or region.

***Pectobacterium* islands and lineage-specific islands.** Bell et al. [1] described 17 islands in the *Pa* genome absent from other enterobacteria. We examined the extent to which these particular islands are conserved among the three *Pectobacterium* species (Figure 3). We also defined islands that differentiate these three genomes using the Mauve alignment. Supplementary Fig. 1 shows the number of regions greater than 10 kb found in only one or two, but not all three *Pectobacterium* genomes. For the two draft genomes, we allowed islands to span multiple ordered contigs. Frequently, small islands can be gathered into larger “variable regions” that correspond to a likely single evolutionary event,

Figure 3.

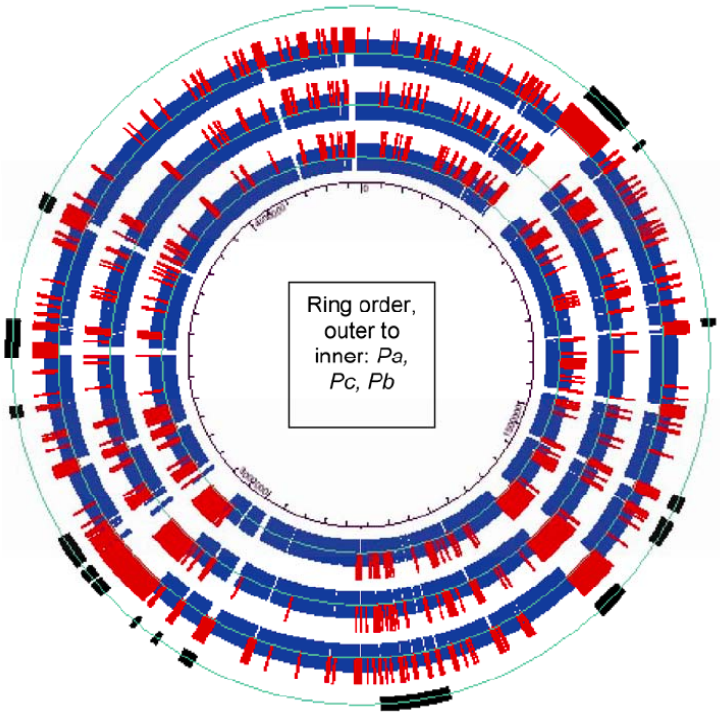


Figure 3. Circular diagram comparing three *Pectobacterium* genomes. This plot is shown using the coordinate system of the complete *Pa* genome. The outer ring illustrates the position of HAI (black boxes) previously reported by Bell et al. 2004. The next ring represents the *Pa* genome with blue corresponding to regions conserved across all three genomes and red corresponding to regions present in only one or two of the three genomes. The next ring represents the *Pc* genome and the inner most ring is the *Pb* genome.

---

such as acquisition of a prophage, and/or a cluster of genes related to a common biological process, such as O-antigen synthesis. Analysis of the functional classes (GO terms) of genes represented in gene islands specific to one or two genomes (see Figure 4) revealed that the predicted function for genes in indels

were in similar proportions to those in the core genome with the exceptions of genes for DNA replication and genes for transcriptional regulation. The abundance of genes for phage in the indels accounts for the overrepresentation of genes for DNA replication. The abundance of unique transcriptional regulatory genes could reflect species adaptation to specific environments.

Figure 4.

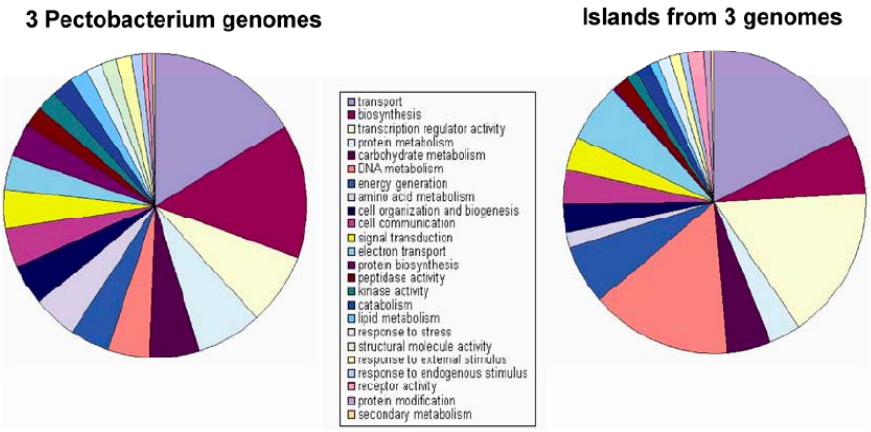


Figure 4. Pie charts comparing Gene Ontology categories between genes found in all Pectobacterium genomes and genes unique to one of the three genomes. These pie charts show the distribution of genes in GO term categories. The chart on the left shows the distribution for all genes in all three *Pectobacterium* genomes and the chart on the right shows the distribution for genes unique in one of the three *Pectobacterium* species. The key in the middle shows the color code for the GO term categories shown in the pie charts.

Manual review of the computational predictions and visual inspection using the Mauve alignment viewer produced Supplementary Figure 1, which describes the 42 variable regions (VR001-VR042). This list includes both regions with a

computational predicted island longer than 10 kb, and clusters of shorter islands that collectively span a region greater than 10 kb in at least one of the *Pectobacterium* species. Genes at the end of each island are listed, as well as putative functions for the genes encoded in the island. The *Pa* horizontally acquired islands identified by Bell et al. [1] are indicated in parentheses. The variable region (VR_000) designations provide an identifier for a polymorphic region flanked by conserved chromosome irrespective of whether or not there are islands (homologous or nonhomologous) at that location. More details about the content and biological significance of these regions are described below. *Pa* ORFs are referred to using the designation given in Bell et al. ([1]) and ORFs in the draft genome sequences are indicated by their ASAP feature ID number. When orthologous ORFs are present in all three genomes, only the *Pa* ORF number is used.

Many previously defined *Pa* HAIs are not conserved in *Pc, Pb* or both. The results presented in the previous section demonstrate that while the HAIs differentiate *Pa* from other closely related animal-associated enterobacterial genomes [1], a substantial number of genes in those regions are not essential determinants of plant pathogenicity, or even the specific soft-rot phenotypes associated with *Pectobacterium* species as a group. Some of the HAIs identified previously in *Pa* are entirely absent from both *Pb* and *Pc*, including HAI2, HAI3, HAI4 and HAI9. Although many of these are phage-related, HAI2 includes genes for a type IV pilus, as well as the *cfa* cluster (PA0511-0614). For other HAIs,

parts are conserved in one or the other new genomes, but not both. Our multiple alignment suggests that several have undergone rearrangements in one or more genomes. Large parts of HAI14, including the nitrogen fixation genes, are present only in *Pa*. Both *Pa* and *Pb* encode metabolic genes and transporters at the same relative chromosomal position, but also a phage suggesting this location may also be a frequent site of insertion of horizontally acquired genes. The HAI7 Type IV system of *Pa* (PA1598-1679) is replaced by two different insertions in *Pc* and *Pb* as expected if this is a hot spot for variation (see VR016 in supplemental Figure 1); however, *Pb* has T4SS (AED 4454-4501) genes, as part of an apparent integrated plasmid at a different site. In contrast HAI5, HAI6, HAI8, HAI15 and HAI 17 are completely or largely conserved among all three genomes. These HAI encode exoplysaccharide and the O antigen (HAI5); a non-ribosomal peptide phytotoxin (HAI6); the Type 3 Secretion System (T3SS) and HecAB agluttinin (HAI8), the *aggA* agglutination adhesion island (HAI15); and a phage insertion (HAI17).

Host range determinants in *Pectobacterium*. Soft rot pathogens are notorious as broad host range pathogens capable of decaying a broad range of monocot and dicot plants. *P. atrosepticum* is notable because it is a well-characterized narrow host range pathogen, found widely on potato and rarely on other solanaceous crops. In other bacterial plant pathogens, such as *Xanthomonas* and *Pseudomonas*, where individual strains can often only infect a few plant hosts, pathogenesis is often restricted by recognition of T3SS effectors by host

plants. Thus, in these systems, in at least some cases, host range limitation is due to host recognition of genes encoded by the pathogen and not because the pathogen is lacking genes required to infect a particular host species.

The draft genome sequences of *Pb* and *Pc* provide some insight into the host range limitations of *Pa*. Unlike *Pseudomonas* and *Xanthomonas* pathogens, the *Pectobacterium* appear to encode few T3SS effector proteins, so it is unlikely that these proteins are limiting the host range of *Pa* to potato. In fact, the *Pa* T3SS is even more restricted than *Pb* and *Pc* (see below). The most striking difference among the three species is that *Pa* encodes genes for production of a putative phytotoxin that are lacking from the other two species (VR004 in supplemental Figure 1). The action or recognition of this toxin by other plant species could restrict the host range of *Pa* in a similar manner as seen with many toxin-producing fungal pathogens. The genomes of the broad host range *Pb* and *Pc* pathogens appear to encode more plant cell wall degrading enzymes, an additional polyketide and peptide synthetase, and several large genes of unknown function, all of which could contribute to pathogenesis on a wide range of plant species. These draft genome sequences do not clearly show if the limitation of *Pa* to potato is due to this pathogen lacking mechanisms required for pathogenesis on other species or due to *Pa* inducing defenses of other species, but the draft genomes do provide a number of high priority targets for answering this fundamental question about soft rot pathogenesis.

Plant cell wall degradation

Three *Pectobacterium* genomes encode nearly identical type II secretion systems and similar sets of plant cell wall degrading enzymes. *Pectobacterium* wilt and rot symptoms are caused by enzymes, including pectate lyases, polygalacturonases, cellulases, and a rhamnogalacturonase, which degrade the structural components of the plant cell wall [37]. Multiple layers of regulation control synthesis of plant cell wall degrading enzymes in soft rot pathogens and the known regulators, including KdgR, ExpRI, RexZ, Crp, and H-NS, are conserved among the three *Pectobacterium*. Yet, there are also important differences in regulators among the three strains, with *Pa* encoding two tandem copies of *pecT* and lacking an AraC-family transcriptional regulator likely to regulate the pectin methyl esterase gene *pmeB*. In addition, PecSM and Pir, key regulators found in *Dickeya* [38, 39] are not present in *Pectobacterium*.

Most of the plant cell wall degrading enzymes, as well as Svx, a protein of unknown function homologous to *X. campestris* AvrXca, are secreted through the type II secretion system (T2SS) [40]. All three *Pectobacterium* species encode homologous T2SS gene clusters in the same locus and all three also encode a pectate lyase and a polygalacturonase adjacent to the T2SS gene cluster. All three species also encode orthologous plant cell wall degrading enzymes including ten pectate lyases, one pectin lyase, four polygalacturonases, two cellulases, and one rhamnogalacturonase (Table 2). The soft rot and stem rot

Table 2. *Pectobacterium* plant cell-wall-degrading enzymes[a]

Enzyme	*Pectobacterium atrosepticum*	*P. brasiliensis*	*P. carotovorum*
Cellulases			
	ABL-0062482	ADT-0000571	AED-0001316
bcsZ	ABL-0064711	ADT-0001191	AED-0002581
Pectate and pectin lyases			
pelA	ECA 4067	AED-0002421	ADT-0004257
pelB	ECA 4068	AED-0003514	ADT-0004256
pelC	ECA 4069	AED-0002430	ADT-0004255
pelI	ECA 1094	AED-0002063	ADT-0001495
pelW	ECA 2402	AED-0001427	ADT-0002894/5
pelX	ECA 4510	AED-0002687	ADT-0001203
pelZ	ECA 4070	AED-0002419	ADT-0002215
	ECA 3112	AED-0001793	ADT-0000798
	ECA 2135	AED-0001225	ADT-0001743
	ECA 2553	AED-0003665	ADT-0003928
	Not present	AED-0002413	Not present
pnl	ECA 1499	AED-0000858	ADT-0001600
Polygalacturonases			
pehA	ECA 1095	AED-0002061	ADT-0000261
pehK	ECA 3552	AED-0003871	ADT-0003117
pehN	ECA 1190	AED-0000675	ADT-0000300
pehX	ECA 3111	AED-0001794	ADT-0000797
peh fragment	ECA 0663	Not present	Not present
Proteases			
	ECA 0879	AED-0000493	ADT-0001419
	ECA 1450	AED-0003719	ADT-0003664
	ECA 2007	AED-0001126	ADT-0001719
	ECA 2785	AED-0004128	ADT-0004008
	ECA 3211	AED-0001839	ADT-0001994
	ECA 2160	AED-0001247	ADT-0001749
	ECA 2074	AED-0003289	ADT-0002858
	ECA 1988	AED-0001062	ADT-0000521
	ECA 1290	AED-0000730	ADT-0001454
	ECA 4192	AED-0002472	ADT-0001132
	ECA 4397	AED-0001980	ADT-0003103
	ECA 2771	AED-0003468	ADT-0003994
	ECA 0388	AED-0000276	ADT-0001318
	ECA 0134	AED-0004146	ADT-0003438
	ECA 0133	AED-0000078	ADT-0001251
	ECA 2802	AED-0001639	ADT-0000726
	ECA 2964	AED-0001716	Not present
	Not present	AED-0003397	Not present
	Not present	AED-0004092	Not present
	ECA 1004	Not present	ADT-0003580
	Not present	AED-0003242	ADT-0002516
	Not present	Not present	ADT-0004360
	Not present	AED-2871	ADT-3651
	ECA 0980	Not present	Not present
	Not present	AED-0004079	Not present
	Not present	Not present	ADT-2640
	Not present	Not present	ADT-3526 and ADT-2470

[a] *Pectobacterium* protease homologs not also present in *Escherichia coli* are included as potential plant cell-wall-degrading enzymes.

symptoms caused by these three species are similar and the conservation of the plant cell wall enzymes is likely to account for much of the similarity in symptoms. Both *Pc* and *Pb* encode putative cell wall degrading enzymes that are not present in *Pa*. For example, both *Pc* and *Pb* have an indel that consists of a GntR regulatory protein homolog (ADT-0003398; AED-0003444), a putative permease that could import the digested polymer (ADT-0002526; AED-0001909),

a gene encoding a putative polysaccharide deacetylase (ADT-0003396; AED-0003442), and a Asp/Glu racemase (ADT-0003397; AED-0003443), which could degrade a host polymer. In addition, *Pc* encodes a cluster of enzyme homologs most closely related to *Clostridium* genes that may also play a role in plant cell wall degradation, including a glycoside hydrolase and a xylan beta-1,4-xylosidase (ADT-0002737-8). Pectolytic *Clostridium* species are often present in decaying root and tuber crops along with *Pectobacterium*, thus it is not surprising that there would be evidence of horizontal gene transfer between these genera. In addition to the two genes mentioned above, 11 additional *Pc* genes and two *Pb* genes are most similar to *Clostridium* genes.

Metalloproteases contribute to *Pectobacterium* virulence and multiple *Pectobacterium* gene islands encode novel proteases. *Pectobacterium* proteases also contribute to plant cell wall degradation [41] (see Supplemental Table 1) and numerous proteases not found in related animal pathogens are present in *Pectobacterium*. Those proteases that have been experimentally examined are secreted via a type 1 secretion system (T1SS) [42, 43]. Many of the *Pectobacterium* proteases are also found in *Dickeya* and a few are present in *P. syringae*, suggesting that their importance in virulence may be under-estimated. Plants defend against proteases by production of enzymes inhibitors [44]; the presence of numerous proteases in gene islands and the likely co-evolution of microbial peptidases and plant inhibitors are reminiscent of the co-evolution of type III effector proteins and plant disease resistance genes.

Suppression of Plant Defenses – the T3SS and phytotoxins.

All three species encode homologous type III secretion systems, but P. atrosepticum lacks hrpK.. In contrast to hemibiotrophic plant pathogens such as *Pseudomonas syringae* and *X. campestris*, the T3SS of pectobacteria does not appear to be essential for growth on potato, since *Pw* does not encode a T3SS but can cause disease on potato (Kim et al., in prep). However the T3SS does contribute to virulence of other *Pectobacterium* species [45, 46]. Mutation of genes required for the secretion apparatus or conserved effector proteins from *Pa* 1039 leads to a reduction in virulence on potato [46]. The T3SS of the three sequenced *Pectobacterium* species are homologous and in the same locus. All genes known to be required for functional T3SS are present in all three species, however, only two of the three, *Pc* and *Pb*, elicit a type III-dependent hypersensitive response [17] on tobacco plants, which indicates that resistance genes in tobacco recognize and respond to type III effector proteins delivered from *Pectobacterium*.

We examined the three genomes for differences in putative T3SS effectors to account for this difference. All three *Pectobacterium* species encode homologs to T3SS secreted proteins identified in *Pa*. Both *Pc* and *Pb*, but not *Pa*, encode HrpK, a T3SS-secreted protein that aids in translocation of effectors across the plant cell wall in *P. syringae* [47]. Mutation of *hrpK* in *Pc* WPP14 did not eliminate the ability of this strain to elicit the HR and expression of the *Pc* WPP14 *hrpK*

gene from a plasmid in *Pa* did not confer HR elicitation onto *Pa* (not shown). Thus, the lack of HR elicitation by *Pa* SCRI1043 is not due to lack of *hrpK*.

Other secretion systems – T4SS, T5SS

Plant pathogenic bacteria encode numerous secretion systems that contribute to virulence in addition to the T1SS, T2SS, and T3SS. The T4SS, which is required for plasmid conjugation and which, like the T3SS, is capable of secreting proteins to the extracelllular milieu and of translocating proteins into host cells, is present in different locations in *Pa* (HAI17; VR016 Supplementary Figure 1) and *Pb* (VR028 AED4454-4501). A remnant of a T4SS is present in *Pc* (ADT-0003337). Currently, it is not possible to predict which proteins travel the T4SS, thus, although it appears that the *Pa* T4SS may contribute to virulence [1], whether or not proteins secreted via the T4SS are conserved among *Pa* and *Pb* remains unknown.

The type five secretion system (T5SS), which includes auto-transporter and two-partner secretion, is the simplest of the secretion systems [48]. T5SS play important roles in pathogenicity of many bacterial pathogens, including *Dickeya*, where the T5-secreted HecA hemagglutinin promotes attachment to leaf surfaces [49]. Several large proteins encoded in *Pectobacterium* indels are likely to be secreted via T5SS, including serine protease, hemolysin, and hemagglutinin homologs.

Polyketide and peptide synthetases

Secondary metabolites produced by polyketide and/or peptide sythetases are important fitness and virulence factors in *P. syringae*, where these exceptionally large proteins produce toxins active against plants and microbes as well as siderophores critical for obtaining iron. The *cfa* locus (HAI2) of *Pa* is absent from *Pb* and *Pc*. A *Pa* region in HAI6 encodes genes (PA1487-1488) similar to the pore-forming phytotoxin syringomycin synthetase from *P. syringae* [50] and is conserved among all three genomes, but fragmented in the two draft genomes. In addition, *Pc* and *Pb* encode a polyketide or peptide synthetase system not found in *Pa* (VR006; ADT-0002503-2509), which is most similar to ones from distantly related Gram-positive bacteria and cyanobacteria, including several that produce toxins. This variability in secondary metabolite production capabilities suggests differences in how the *Pectobacterium* species interact with either their plant hosts or competing microbes, and calls for further characterization of the synthesis products, analysis of the distribution of these systems among *Pectobacterium* isolates, and cross-species growth inhibition assays.

Motility and Taxis

The *Pectobacterium* species are all capable of swimming and swarming and the entire flagellar gene cluster is encoded in one locus, similar to that of *D. dadantii* and *Yersinia*, but differing from *E. coli* and *S. enterica*. *Pectobacterium* encodes two flagellin homologs. In all three *Pectobacterium*, one of the flagellin genes is located in the midst of the flagella gene cluster while the second flagellin is

encoded elsewhere in a locus that contains numerous indels and rearrangments. All three flagellin genes encode a flg22 peptide sequence, suggesting that *Pectobacterium* could activate Fls2-mediated defenses in Arabidopsis and other host plants able to recognize this particular flagellin peptide.

One hallmark that sets plant pathogens apart from animal pathogens is the large number of taxis proteins encoded in plant pathogens. Related enteric animal pathogens encode five to 12 methyl-accepting chemotaxis (MCPs) receptors, while in contrast, all of the soft rot pathogens encode over 30 MCPs, with *Pa*, *Pc*, and *Pb* encoding 36, 39, and 34 taxis receptors, respectively. Taxis receptors function as heterotrimers of homodimers [51]. Thus, the five receptors commonly found in *E. coli* could form up to 35 different combinations of heterotrimers, although some combinations are more likely than others due to an approximately 10-fold higher concentration of two of the receptors in comparison to the other three. The high number of receptors found in plant pathogens means that thousands of combinations of heterotrimers are possible.

Invertebrate interactions

Many enterobacteria have intimate interactions with insects and since the early part of the 1900s, insects have been suspected of spreading soft rot disease. For example, *Pectobacterium* has been cultured from corn maggot eggs and mouth parts (reviewed in [52]), and transgenic plants with resistance to corn maggots also suffer less from soft rot pathogens. The related pathogen, *D. dadantii* 3937

is notable in that it encodes four *Bacillus thuringiensis* toxin homologs, at least one of which is active against aphids [53]. The genetics of interactions between *Pectobacterium* and invertebrates have only been recently been examined. A single locus, which includes Evf and its regulator, Evr, improves survival of *Pectobacterium* in fruit fly guts [54] [16]. None of the sequenced *Pectobacterium* strains encode either a *Bacillus* toxin homolog or an Evf. *Pa*, *Pb*, and *Pc* all encode the regulator Evr, although in *Pa evr* is a pseudogene. *Pa* may have an additional mechanism for invertebrate interactions since it encodes a cluster of genes homologous to the hemin storage proteins required by *Yersinia* for transmission by fleas [55].

Competition with other microbes – antibiotics, bacteriocins

The *Pectobacterium* species differ in the mechanisms they use to compete with other microbes in the environment. Some strains of *Pectobacterium* produce the carbapenem antibiotic 1-carbapen-2-em-3-carboxylic acid (Car), an indication that *Pectobacterium* is in competition with potato endophytes and secondary invaders for resources [56]. Of the three sequenced strains, the *car* genes are only found in *Pb*. *Pectobacterium* strains also produce multiple forms of carotovoricin, a phage-tail-like bacteriocin which kills closely related strains and species [57, 58]. *Pa* encodes two bacteriocin resistance proteins, which are not homologous to each other, and only one of these two is also found in *Pc* and *Pb*. Both *Pc* and *Pb* encode one gene homologous to the *Yersinia* bacteriocin pesticin (ASAP ID ADT-0002424; AED-0003952) and a second, unrelated

bacteriocin homolog is adjacent to the pesticin homolog in *Pb* (AED-0003953). No bacteriocins are apparent in the *Pa* genome.

Variability in nutrient acquisition and metabolism

Pectobacterium strains vary in the carbon sources they can use [26] and numerous *Pectobacterium* indels appear to be involved in metabolite uptake or degradation, providing a genetic explanation for some of these differences. For example, *Pa* and *Pb*, but not *Pc*, encodes galactonate and gluconate metabolism islands as well as a sucrose isomerase. Thus these *Pa* and *Pb* strains are likely to be able to metabolize more different common plant sugars than *Pc*. This variation suggests that competition among *Pectobacterium* strains goes beyond how well they attack plants and includes which plant nutrients they are able to metabolize.

A protein family unique to the soft rot erwinia

A family of large proteins ranging between 515 and 732 residues was found in all of the soft rot enterobacterial genomes, but is lacking from all other genomes in Genbank. Two members of this protein family are found in *Pa* (ECA1185; ECA1186) and *Pb* (AED-0003594; 2989), while *Pc* encodes three members (AADT-002625; 3321; 4109) and *D. dadantii* 3937 encodes one homolog (ABF-0020188). These proteins each have a conserved F5/8 type C domain that may aid in binding galactose [59]. Other proteins with this domain bind to cell

membranes, thus these proteins could be involved in interactions with the plant cell wall or cell membrane.

CONCLUSIONS

We characterized the pan-genome of *Pectobacterium* by comparing three genomes from strains that are well separated phylogenetically within the genus. Using 454 platform pyrosequencing, to a depth of approximately 20-fold genome coverage allowed assembly into roughly 200 contigs representing approximately 95% of each draft genome. The availability of a high quality assembled genome from a related type strain made it possible for us to align the 200 contigs of each genome into a useful chromosome assembly using the Mauve program. The genes unique to each isolate were identified and aligned making the information most useful for genome comparison available at an affordable cost. However, without completing the genome sequences by closing gaps and resequencing regions of low coverage or ambiguous sequence, it is not possible to rule out the existence of missing or erroneous data, determine the extent of rearrangements among chromosomes, and identify extrachromosomal elements. As the use of 454 and other “next generation” sequencing technologies increases, it is important to weigh the costs and benefits of collecting additional data such as dual- ended sequences of fragments, optical map information, or directed sequences of gaps by standard PCR or cloning methods. Currently, the costs of collecting of this additional data are roughly equivalent to the costs of obtaining

the initial sequence data and warrant consideration when completion of the genome is critical for comparison or interpretation of sequence data. In the case of the *Pectobacterium* genomes presented here, little was known about the pan-genomic content of members of this genus before this report, and analysis of partially completed genomes was sufficient to identify major differences between isolates that will form the basis for further characterization to see if they explain phenotypic variation.

We found that each strain differed in approximately 11-18% of the genome. Regulatory genes were more abundant in the class of uniquely represented genes. This diversity in gene regulatory mechanisms may reflect adaptation to specific ecological niches where each organism must respond to different environmental stimuli such as varying temperatures and alternative stressful conditions. Not surprisingly, these regulators are often clustered with other strain-specific genes, suggesting potential targets of regulation, and many of these variable genes encode products involved in metabolic processes.

A number of genomic differences that we identified might have consequences for virulence of the organisms. We found variation in the content of plant cell wall degrading enzymes, the most obvious players in causing the symptoms of soft-rot disease, as well as diversity in the content of genes encoding metalloproteases, which are also important for degradation of host cell walls. Each strain has unique genes for production of phytoxins, another likely

determinant of host-range. Differences in the secretion abilities of these organisms are suggested by variation in the presence and location of type IV (and other) secretion systems among the genomes. Each genome had a large but not fully overlapping collection of genes for motility and chemotaxis. Genes for surface proteins such as O-antigens, pili and adhesins showed significant variability among genomes for factors that could either suppress or trigger host defenses. A large fraction of genes that differ between these organisms are of unknown function, including a family of proteins that appears unique to the soft rot pathogens. Figuring out which of the factors identified in this study are key determinants of the phenotypic differences between these pathogens will require sequencing of additional genomes and experimental characterization of gene functions, particularly sensitive assays for even subtle roles in virulence. The *Pa*, *Pb* and *Pc* isolates examined in this report represent relative extremes with respect the phylogenetic diversity within this genus, and their genomes vary considerably in genomic content. Resolving which changes are responsible for niche specialization is challenging given the magnitude of genomic differences. Sequences of a few additional members of the genus, particularly isolates more closely related to each of the groups already sampled or isolates with similar phenotypes would likely help pinpoint a smaller number of important factors.

LITERATURE CITED

1. Bell, K.S., et al., *Genome sequence of the enterobacterial phytopathogen Erwinia carotovora subsp atroseptica and characterization of virulence factors.* Proc. Natl. Acad. Sci. U S A, 2004. **101**(30): p. 11105-11110.

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

3. Feil, H., et al., *Comparison of the complete genome sequences of Pseudomonas syringae pv. syringae B728a and pv. tomato DC3000.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(31): p. 11064-11069.

4. Joardar, V., et al., *Whole-genome sequence analysis of Pseudomonas syringae pv. phaseolicola 1448A reveals divergence among pathovars in genes involved in virulence and transposition.* Journal of Bacteriology, 2005. **187**(18): p. 6488-6498.

5. Thieme, F., et al., *Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium Xanthomonas campestris pv. vesicatoria revealed by the complete genome sequence.* Journal of Bacteriology, 2005. **187**(21): p. 7254-7266.

6. Jones, J.D.G. and J.L. Dangl, *The plant immune system.* Nature 2006. **444**: p. 323-329.

7. Welch, R.A., et al., *Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli.* Proc Natl Acad Sci U S A, 2002. **99**(26): p. 17020-4.

8. Tettelin, H., et al., *Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial "pan-genome".* Proc Natl Acad Sci U S A, 2005. **102**(39): p. 13950-5.

9. Kettler, G.C., et al., *Patterns and Implications of Gene Gain and Loss in the Evolution of Prochlorococcus.* PLoS Genet, 2007. **3**(12): p. e231.

10. Willenbrock, H., et al., *Characterization of probiotic Escherichia coli isolates with a novel pan-genome microarray.* Genome Biol, 2007. **8**(12): p. R267.

11. McCarter-Zorner, N.J., et al., *Soft rot Erwinia in surface and underground waters in Southern Scotland and in Colorado, United States.* J. Appl. Bacteriol., 1984. **57**: p. 95-105.

12. McCarter-Zorner, N.J., et al., *Soft rot Erwinia bacteria in the rhizosphere of weeds and crop plants in Colorado, United States and Scotland.* J. Appl. Bacteriol., 1985. **59**: p. 357-368.

13. Harrison, M., et al., *Waste potato dumps as sources of insects contaminated with soft rot coliform bacteria in relation to re-contamination of pathogen-free potato stocks.* Potato Research, 1977. **20**(1): p. 37-52.

14. Molina, J.M., D. Harrison, and J.W. Brewer, *Transmission of Erwinia carotovora var. atroseptica by Drosophila melanogaster Meig. I. Acquisition and transmission of the bacterium.* Am. Potato J., 1974. **51**: p. 245-250.

15. Phillips, J.A. and A. Kelman, *Direct fluorescent antibody stain procedure applied to insect transmission of Erwinia carotovora.* Phytopathology, 1982. **72**(7): p. 898-901.

16. Basset, A., et al., *The phytopathogenic bacteria Erwinia carotovora infects Drosophila and activates an immune response.* Proc. Natl. Acad. Sci. USA, 2000. **97**(7): p. 3376-3381.

17. Gardan, L., et al., *Elevation of three subspecies of Pectobacterium carotovorum to species level: Pectobacterium atrosepticum sp nov., Pectobacterium betavasculorum sp nov and Pectobacterium wasabiae sp nov.* Int. J. Syst. Evol. Microbiol., 2003. **53**: p. 381-391.

18. Duarte, V., et al., *Characterization of atypical Erwinia carotovora strains causing blackleg of potato in Brazil.* J Appl Microbiol, 2004. **96**(3): p. 535-45.

19. Ma, B., et al., *The host range and molecular phylogenies of the soft rot enterobacterial genera Pectobacterium and Dickeya.* Phytopathology, 2007. **97**: p. 1150-1163.

20. Barnard, A.M. and G.P. Salmond, *Quorum sensing in Erwinia species.* Anal Bioanal Chem, 2007. **387**(2): p. 415-23.

21. Alfano, J.R., et al., *The Pseudomonas syringae Hrp pathogenicity island has a tripartite mosaic structure composed of a cluster of type III secretion genes bounded by exchangeable effector and conserved effector loci that contribute to parasitic fitness and pathogenicity in plants.* Proc Natl Acad Sci U S A, 2000. **97**(9): p. 4856-61.

22. Brito, B., et al., *A signal transfer system through three compartments transduces the plant cell contact-dependent signal controlling Ralstonia solanacearum hrp genes.* Mol Plant Microbe Interact, 2002. **15**(2): p. 109-19.

23. Quevillon, E.S., V.; Pillai, S.; Harte, N.; Mulder, N.; Apweiler, R.; Lopez, R., *InterProScan: protein domains identifier.* Nucleic Acids Res., 2005. **33**: p. W116-120.

24. Cui, Y., et al., *Identification of a global repressor gene, rsmA, of Erwinia carotovora subsp. carotovora that controls extracellular enzymes, N-(3-oxohexanoyl)-L-homoserine lactone, and pathogenicity in soft-rotting Erwinia spp.* J Bacteriol, 1995. **177**(17): p. 5108-15.

25. Triplett, L.R., Y. Zhao, and G.W. Sundin, *Genetic differences between blight-causing Erwinia species with differing host specificities, identified by suppression subtractive hybridization.* Appl Environ Microbiol, 2006. **72**(11): p. 7359-64.

26. Yap, M.-N., J.D. Barak, and A.O. Charkowski, *Genomic diversity of Erwinia carotovora subsp. carotovora and its correlation with virulence.* Appl. Envir. Microbiol., 2004. **70**: p. 3013-3023.

27. Sambrook, J. and D.W. Russell, *Molecular Cloning. A Laboratory Manual.* 3rd ed. 2001, Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

28. van Hijum, S.A.F.T., et al., *Projector 2: contig mapping for efficient gap-closure of prokaryotic genome sequence assemblies.* Nucl. Acid. Res., 2005. **33**: p. W560-W566.

29. Darling, A.C.M., B.; Blattner, F.R.; Perna, N.T., *Mauve: multiple alignment of conserved genomic sequence with rearrangements.* Genome Res., 2004. **14**: p. 1394-1403.

30. Altschul, S.F., et al., *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Res., 1997. **25**(17): p. 3389-3402.

31. Delcher, A.L., et al., *Improved microbial gene identification with GLIMMER.* Nucleic Acids Res, 1999. **27**(23): p. 4636-41.

32. Glasner, J.D., et al., *ASAP: a resource for annotating, curating, comparing, and disseminating genomic data.* Nucleic Acids Res, 2006a. **34**: p. D41-45.

33. Bogdanove, A.J., D.W. Bauer, and S.V. Beer, *Erwinia amylovora secretes DspE, a pathogenicity factor and functional AvrE homolog, through the Hrp (type III secretion) pathway.* J. Bacteriol., 1998. **180**(8): p. 2244-7.

34. Ried, J.L. and A. Collmer, *An nptI-sacB-sacR cartridge for constructing directed, unmarked mutations in gram-negative bacteria by marker-exchange-eviction mutagenesis.* Gene, 1987. **57**: p. 239-246.

35. Margulies, M., et al., *Genome sequencing in microfabricated high-density picolitre reactors.* Nature, 2005. **437**: p. 376-380.

36. Glasner, J.D., et al., *ASAP: a resource for annotating, curating, comparing, and disseminating genomic data.* Nucleic Acids Res, 2006. **34**(Database issue): p. D41-5.

37. Toth, I.K. and P.R.J. Birch, *Rotting softly and stealthily.* Curr. Opin. Plant Biol., 2005. **8**: p. 424-429.

38. Praillet, T., S. Reverchon, and W. Nasser, *Mutual control of the PecS/PecM couple, two proteins regulating virulence-factor synthesis in Erwinia chrysanthemi.* Mol Microbiol, 1997. **24**(4): p. 803-14.

39. Nomura, K., et al., *The pir gene of Erwinia chrysanthemi EC16 regulates hyperinduction of pectate lyase virulence genes in response to plant signals.* Proc Natl Acad Sci U S A, 1998. **95**(24): p. 14034-9.

40. Corbett, M., et al., *Identification of a new quorum-sensing-controlled virulence factor in Erwinia carotovora subsp. atroseptica secreted via the type II targeting pathway.* Mol Plant Microbe Interact, 2005. **18**(4): p. 334-42.

41. Marits, R., et al., *Isolation of an extracellular protease gene of Erwinia carotovora subsp carotovora strain SCC3193 by transposon mutagenesis and the role of protease in phytopathogenicity.* Microbiol., 1999. **145**: p. 1959-1966.

42. Delepelaire, P. and C. Wandersman, *Characterization, localization and transmembrane organization of the three proteins PrtD, PrtE and PrtF necessary for protease secretion by the gram-negative bacterium Erwinia chrysanthemi.* Mol. Microbiol., 1991. **5**(10): p. 2427-2434.

43. Wandersman, C., P. Delepelaire, and S. Letoffe, *Secretion processing and activation of Erwinia chrysanthemi proteases.* Biochimie, 1990. **72**(2-3): p. 143-6.

44. Quilis, J., et al., *A potato carboxypeptidase inhibitor gene provides pathogen resistance in transgenic rice.* Plant Biotechnol. J., 2007. **5**(4): p. 537-553.

45. Rantakari, A., et al., *Type III secretion contributes to the pathogenesis of the soft- rot pathogen Erwinia carotovora: partial characterization of the hrp gene cluster.* Mol. Plant-Microbe Interact, 2001. **14**(8): p. 962-968.

46. Holeva, M.C., et al., *Use of a pooled transposon mutation grid to demonstrate roles in disease development for Erwinia carotovora subsp atroseptica putative type III secreted effector (DspE/A) and helper (HrpN) proteins.* Mol. Plant-Microbe Interact., 2004. **17**(9): p. 943-950.

47. Petnicki-Ocwieja, T., K. van Dijk, and J.R. Alfano, *The hrpK operon of Pseudomonas syringae pv. tomato DC3000 encodes two proteins secreted by the type III (Hrp) protein secretion system: HopB1 and HrpK, a putative type III translocator.* J. Bacteriol., 2005. **187**(2): p. 649-663.

48. Henderson, I.R., et al., *Type V protein secretion pathway: the autotransporter story.* Microbiology and Molecular Biology Reviews, 2004. **68**(4): p. 692-+.

49. Rojas, C.M., et al., *HecA is a member of a class of adhesins produced by diverse pathogenic bacteria and contributes to the attachment, aggregation, epidermal cell killing, and virulence phenotypes of Erwinia chrysanthemi EC16 on Nicotiana clevelandii seedlings.* Proc. Natl. Acad. Sci. U S A, 2002. **99**: p. 13142-13147.

50. Bender, C.L., F. Alarcon-Chaidez, and D.C. Gross, *Pseudomonas syringae phytotoxins: mode of action, regulation, and biosynthesis by peptide and polyketide synthetases.* Microbiol Mol Biol Rev, 1999. **63**(2): p. 266-92.

51. Parkinson, J.H., P. Ames, and C.A. Studdert, *Collaborative signaling by bacterial chemoreceptors.* Curr. Opin. Microbiol., 2005. **8**: p. 116-121.

52. Gnanamanickam, S.S., *Plant-associated bacteria*. 2006, Dordrecht: Springer. vi, 712 p.

53. Grenier, A.M., et al., *The phytopathogen Dickeya dadantii (Erwinia chrysanthemi 3937) is a pathogen of the pea aphid.* Appl. Environ. Microbiol., 2006. **72**: p. 1956-1965.

54. Muniz, C.A., et al., *Erwinia carotovora Evf antagonizes the elimination of bacteria in the gut of Drosophila larvae.* Cellular Microbiology, 2007. **9**(1): p. 106-119.

55. Hinnebusch, B.J., R.D. Perry, and T.G. Schwan, *Role of the Yersinia pestis hemin storage (hms) locus in the transmission of plague by fleas. .* Science, 1996. **273**: p. 367-370.

56. Parker, W.L., et al., *SQ 27,860, a simple carbapenem produced by species of Serratia and Erwinia.* J Antibiot (Tokyo), 1982. **35**(6): p. 653-60.

57. Nguyen, H.A., J. Kaneko, and Y. Kamio, *Temperature-dependent production of carotovoricin Er and pectin lyase in phytopathogenic Erwinia carotovora subsp carotovora Er.* Bioscience Biotechnology and Biochemistry, 2002. **66**(2): p. 444-447.

58. Nguyen, H.A., et al., *DNA inversion in the tail fiber gene alters the host range specificity of carotovoricin Er, a phage-tail-like bacteriocin of phytopathogenic Erwinia carotovora subsp carotovora Er.* Journal of Bacteriology, 2001. **183**(21): p. 6274-6281.

59. Newstead, S.L., et al., *Galactose recognition by the carbohydrate-binding module of a bacterial sialidase.* Acta Crystallogr. D. Biol. Crystallogr., 2005. **61**: p. 1483-1491.

CHAPTER 3: Transcription profiling during multiple routes of potato infection reveal a tissue-specific role for the T3SS in *Pectobacterium carotovorum*

ABSTRACT:

In this chapter, we present data that will be published in a comparative study of the expression patterns of *P. carotovorum* during multiple conditions that trigger virulence. We look at transcriptional profiles during tuber, stem, and leaf infections, and when the T3SS is artificially triggered. We provide data that suggests the T3SS and associated genes are required for virulence in the leaf, but not stem or tuber infections. Further, we show through confocal microscopy and qRT data that the genes of the T3SS are expressed during the first 7 hours of leaf infection. In order to understand the significance of a T3SS during leaf infection, we used a promoter-trap screen (Differential Fluorescence Induction screen) and DNA microarrays to determine the regulon of HrpL, the alternative sigma factor responsible for regulation of the T3SS cluster and known effectors. We show that despite the requirement for the T3SS during leaf infections, only 8 genes, in addition to the T3SS apparatus genes are co-regulated with T3SS. Lastly, we use DNA arrays to determine the transcriptional profiles during stem and tuber infections to understand key virulence factors associated with those infection routes, which do not require a T3SS. The data suggests that the T3SS is not delivering a large suite of effector genes and is a leaf- specific virulence factor.

INTRODUCTION:

P. carotovorum is a necrotrophic pathogen that causes soft rot disease on plant species from over 24 orders of angiosperms. It is a ubiquitous pathogen that survives in soil, both surface and ground water, and has been associated with a variety of invertebrates [1-4]. Soft rot diseases that develop from *P. carotovorum* infections lead to economically significant losses of crops such as carrot, tobacco, and potato. *P. carotovorum* causes soft rot disease, in part, by secreting over 16 plant macerating enzymes that target plant cell sugars such as polygalacturonase [5-7]. These enzymes degrade plant cell walls and cause the characteristic rotting seen on infected plants. In addition to macerating enzymes, *P. carotovorum* possess adhesions [8], antimicrobial peptides [9], metalloproteases[10], and a T3SS that it may use to infect a host and establish disease [11].

T3SS are used by Gram-negative bacterial pathogens of both animals and plants [12]. The T3SS translocates virulence proteins, referred to as effector proteins, into host cells where they interact and interfere with components of host defense responses (reviewed in [13, 14]). Delivered effectors may inhibit basal defense responses, which can be triggered by microbe-associated molecular patters (MAMPs) such as flagellin and LPS [15]. T3SS effectors thus can subvert defenses, allowing the pathogen to persist. Plant hosts, however, have evolved to use the presence of effectors as a means of recognizing infection and thus re-establishing defense responses [16]. Plant Resistance proteins (R

proteins) can directly or indirectly sense a delivered effector and re-instate a defense response, referred to as a hypersensitive response (HR), which leads to programmed cell death and inhibition of bacterial growth [17]. Pathogens such as *P. syringae* strains may possess a suite of effectors. Effector suites can include upwards of 30 genes, individuals of which can be gained by horizontal acquisition events or just as easily lost by selective pressure in a host with a corresponding R protein [18]. Thus, phytopathogens and their hosts are often described as being locked in an evolutionary arms race [19].

The T3SS itself is assembled from proteins which retain structural similarity to bacterial flagellum proteins (reviewed in [20]). The genes encoding the T3SS typically lie in clusters of structural and regulatory genes called hrp (HR and Pathogenicity) clusters and they tend to be flanked by genes encoding secreted effectors and their chaperones [21]. The structural genes of the T3SS are highly conserved in diverse species, but the regulators linked within the complexes are variable (reviewed in [22]). Phylogenetic comparisons suggest that T3SS have been acquired by horizontal transmission in diverse plant and animal pathogens at independent stages in their evolutionary histories [19, 23]. Two distinct clades of T3SS emerge in model phytopathogens; Group I in *Pectobacterium* and *Pseudomonas* species and Group II T3SS in *Ralstonia* and *Xanthomonas* species [24]. In *P. carotovorum* and *P. syringae,* an alternative sigma factor, *hrpL,* is the main regulator of genes in the hrp cluster [21, 25], which is likely due to linkage to the T3SS encoded on similar pathogenicity islands that were incorporated into the ancestors of these bacteria independently.

In *P. syringae,* HrpL is known to regulate not only the hrp cluster, but all effector-encoding genes. In *P. carotovorum,* HrpL is also known to regulate the hrp cluster and the one identified effector-encoding gene, *dspE.*

The role of the T3SS in *P. carotovorum* pathogenesis is poorly defined. Unlike the generally biotrophic pathogen *P. syringae, P. carotovorum* does not need to continually subvert host defenses since infection results in death of the host by macerating enzymes. It has been suggested that *P. carotovorum* uses its T3SS early during infection to suppress defense responses prior to generating a sufficient amount of macerating enzymes to establish disease on the host [26].

In this chapter, we address the contribution of the T3SS to *P. carotovorum* during different infection routes and define the regulon of HrpL to identify potential additional effectors. Our findings suggest that the T3SS in *P. carotovorum* is expressed and required for virulence during leaf infection, but HrpL regulates a small number of genes outside the hrp cluster.

METHODS:

Bacterial strains, plasmids and media. Bacterial strains were maintained in Lysogeny Broth (LB) agar or 2xYT broth, containing appropriate antibiotics. The compositions of LB medium, M9 Minimal Media, 2xYT media are published in Sambrook and Russell [27]. The Minimal Media for *hrpL* repression was described in Chang et al. 2005 [28]. When required, antibiotics and drugs were supplemented at the following concentrations: ampicillin (Amp), 100 ug/mL;

chloramphenicol (Cm), 30 ug/mL; gentamycin (Gm), 25 ug/mL; kanamycin (Km), 30 ug/mL; rifampicin (Rif), 100 ug/mL; spectinomycin (Spec), 50 ug/mL; and tetracycline (Tet), 5 ug/mL. Media was solidified with 1.5% BD Bacto-agar. Cultures were grown at 28ºC and, if in liquid, shaken at 250 r.p.m.)

DNA manipulations. Standard procedures were used for: plasmid and chromosomal DNA isolation; electroporation; restriction endonuclease digestions; ligations; gel electrophoresis; and tri-parental mating. Enzymes were obtained from either New England Biolabs or Invitrogen. Nucleotide sequences were determined by the UNC-CH Genome Analysis Facility and sequences were analyzed using BLAST.

Virulence assays on potato and *A. thaliana* leaves. Cultures were grown over night in rich media, washed twice in 10 mM MgCl2, and re-suspended at 10^4 CFU/mL in 10 mM MgCl2. Cultures were infiltrated into the leaves of either 4-5 week old Yukon gold potatoes or 4-5 week old *A. thaliana* using a needle-less syringe. At subsequent time points, infected leaves were cored with a gauge 3 (6 mm diameter) corer and the tissue ground in the presence of 10 mM MgCl2. The CFU/mL of present bacteria was quantified by titration and plating on LB plates with the appropriate antibiotics. Each experiment contained 4 internal replicates.

Plasmids. pBAD::hrpLWPP14 construction: hrpLWPP14 (*hrpL*) was PCR amplified by using Pfu (Stratagene) then cloned into pCFS40 as described in

Chang et al. 2005 [28]. pBAD::hrpL exhibited tight regulation in the absence of arabinose and high expression in the presence of arabinose in minimal media for native *hrpL* repression (data not shown). 200 mM of arabinose was optimal for *hprL* induction. Differential Fluorescence Induction (DFI) Vectors: Construction was described in Chang et al. 2005 [28]. Briefly, pBBR1-MCS2 was modified to carry RNA terminators, and mutant GFP3, which has an excitation peak corresponding to the argon laser (488 nM) of the FACS. Three stop codons in each frame were annealed and cloned upstream from the Shine-Delgarno sequence as an EcoRI-XbaI fragment, yielding vector 125.1. PCR products were directly cloned into EcoRI-BamHI digested 125.1 vector following the sticky-end PCR protocol, resulting in DFI vectors.

Library Construction for Promoter-Trap Screen. DNA was extracted from *P carotovorum* WPP14, purified by using Epicentre MasterPure DNA extraction kit, and partially digested with either Tsp509I, or AluI, BstUI, HaeIII, and RsaI. Fragments from 1–1.6 kb, 1.6–3 kb, and 3–4 kb were extracted and cloned into either EcoRI- or SmaI-digested and shrimp alkaline phosphatase (SAP)-treated DFI vectors. *E. coli* colonies carrying approximately 66,000 library fragments were pooled and mated en masse by modified triparental mating with WPP14 + pBAD::hrpL and pRK2013.

The DFI Screen. The screen was done as described in Chang et al. 2005 [28]. Briefly, we screened the DFI library under the inducing conditions of 200 mM

arabinose for 22 h before FACS screening. FACS was performed on a MoFlo (Cytomation) and analysis was performed on a FACscan from Becton Dickinson. One mL of culture was diluted into 400 µl of 1× PBS. HrpL-inducible gene fragments were identified with 4 subsequent sorts. The first, in the absence of induction, the least fluorescent cells (30%) were collected. These cells were grown in inducing conditions and a small population (less than 2% of total population) of highly-fluorescent cells was collected; these two sorts were repeated, except final fluorescent cells were collected in individual wells of a 364 well plate. Individual library clones were grown in these plates, their library fragments were amplified from the DFI vector, and amplicons were sequenced.

Sequencing library clones: Approximately 2600 clones were sequenced. Candidate HrpL-induced gene fragments were amplified from cells harboring the DFI plasmids using Taq polymerase and primers, HYZ163 and HYZ166 [28]. PCR products were treated with 5 units of exonuclease I and 0.5 unit of SAP at 37°C for 40 min and heat terminated at 80°C for 30 min. Sequences were aligned to the WPP14 genome sequence.

Construction of the *P. carotovorum* mutants. To delete *hrpL* we used splicing over-lap extension (SOE) PCR to create a construct with the Cm cassette flanked by 1 kb of sequence surrounding *hrpL* on each side. Briefly, a 1 kb region upstream of *hrpL* was amplified using SOE primers containing a RE site for the insertion of a Cm cassette on the 3' end. A 1 kb downstream region was

amplified similarly with the RE site on the 5' end with SOE primers. These amplicons were then fused together in a two-step SOE PCR reaction and the SOE product was cloned into pCR2.1 Topo TA cloning vector (ampR) (Invitrogen) to make pCR2.1SOE*hrpL*, which was then digested with RE. The Cm cassette was amplified from pKD3 using modified primers from [29] to contain RE sites on both ends. The Cm cassette was then digested with RE and ligated to the linear pCR2.1SOEhrpL to make pCR2.1ΔhrpL. This plasmid was introduced into electro-competent *P. carotovorum* WPP14 by electroporation. The resulting WPP14: pCR2.1Δ*hrpL* strain was then grown in potassium phosphate buffer supplemented with Cm, but not Amp. After over-night growth, the cultures were transferred to 2xYT supplemented with Cm and passaged for 3 days, at which point the cultures were replica plated on LB plates that contained either Cm and Amp or just Cm to identify strains that had lost the plasmid but undergone double-homologous recombination for marker exchange. The resulting deletion-insertion mutants were verified by PCR analysis and sequencing.

***A. thaliana* Growth Conditions.** *A. thaliana* were grown in under short-day lighting conditions in a Conviron Growth Chamber at 23ºC.

Potato growth conditions and inoculation methods. Certified seed potatoes (*Solanum tuberosum* subsp. *tuberosum* var. Superior) were grown in a growth chamber at 26 ˚C with a 12 hour photoperiod. Stems of 4 weeks-old plants were

inoculated by stabbing a needle and injecting a bacterial suspension of WPP14 (0.3 OD_{600}, which is $1.5x10^8$ CFU/ml) [30, 31]. To prepare the bacterial suspensions, cells from 24 hour-old bacterial cultures grown on LB agar medium incubated at 37˚C were suspended in water. Each potato stem was inoculated in five different sites approximately 3 cm from each other. After inoculation, plants were placed into plastic bags and incubated at 26˚C for 13.5 hours, at which time the bacteria are in log phase growth in stems. For tuber inoculation, potato tubers (*Solanum tuberosum* subsp. *tuberosum* var. Superior) were surface sanitized by soaking them in 0.5% bleach for 10 minutes. Tubers were rinsed with distilled water and allowed to dry at room temperature for 24 hours. Tubers were inoculated with bacterial suspension of *P. carotovorum* (0.3 OD_{600}) using 10 µl sterile tips. There were at least 15 inoculation sites per tuber. Tubers were then placed into plastic bags and incubated at 26˚C for 10 hours, at which time the bacteria are in log phase growth in tubers. For leaf infections, cultures of bacteria at OD_{600} =1.0 were used to infiltrate leaves. After given time points leaves were harvested for RNA extraction.

Bacterial sample collection from plants and RNA isolation. For bacterial cells collection from stems and tubers, approximately 90 and 75 inoculations sites, respectively, were sampled. Plant material was placed in 3 cooled sterile RNAse-free mortars pre-filled with 17.5 ml of DEPC water and 1.25 ml of ice cold EtOH/Phenol stop solution (5% water-saturated phenol pH<7.0) to stabilize cellular RNA and stop degradation. Plant samples were gently ground with a

sterile RNAse-free pestle. Supernatants were collected in two 15 ml falcon tubes and a modified hot SDS/hot phenol method [35] was followed. Briefly, bacterial supernatants were harvested at 8,000 rpm for 2 minutes at 4 °C. The supernatant was aspirated and the pellet frozen in dry ice-alcohol mix. Pellets were resuspended in 1.6 ml of lysis buffer (TE: 10 mM Tris - 1 mM EDTA, pH 8.0) and 0.5 mg/ml Ready-Lyse™ Lysozyme Solution (Epicentre Biotechnologies, Madison, WI) and 160 µl of 10% SDS solution, mixed and incubated at 65 ˚C for 2 minutes. Then, 176 µl of 1 M NaOAc (pH 5.2) was added and mixed with the lysate. Equal volumes of lysate were transferred to RNase-free 2 ml microfuge tubes and mixed by inversion with water saturated phenol. Tubes were incubated in 65 °C waterbath for 6 minutes. Then, tubes were placed on ice to chill for 2 minutes and centrifuged at 14,000 rpm for 10 min at 4 °C. The aqueous layer was then transferred to a new 2 ml microfuge tube containing an equal volume of chloroform. Tubes were mixed by inversion, and then centrifuged at 14,000 rpm for 10 minutes at 4 °C. The aqueous layer was split into equal volumes in 1.5 ml Eppendorf tubes and ethanol precipitated by adding 1/10 volume of 3M NaOAc (pH 5.2), 1 mM EDTA and 2 volumes of cold EtOH to each sample. Samples were mixed and placed at -80 0C overnight. RNA was pelleted by centrifugation at 14,000 rpm for 30 minutes at 4 0C. Then, pellets were washed with ice cold 80% EtOH and spun down at 14, 000 rpm for 10 minutes. Supernatant was removed and pellets were air dried for 30 minutes in the fume hood. Pellets were resuspended with DEPC-treated water. All resuspended samples were pooled together (100 µl) and DNA contamination

was removed using Turbo DNA-free (Ambion, Inc.) with the rigorous protocol and 5 µl of enzyme. Additional cleanup of the sample was performed using RNeasy Mini Kit (Qiagen, Valencia, CA). Purity and concentration of the RNA was verified with a spectrophotometric analysis (NanoDrop ND-1000; NanoDrop technologies, Wilmington, DE). RNA integrity was measured using an Agilent 2100 Bioanalyzer (Agilent Technologies, Inc.). RNA was stored at -80 0C until use.

Transcription profile conditions: pBAD: hrpL was moved into WPP14Δ*hrpL* via trip-parental mating. WPP14Δ*hrpL* and WPP14Δ*hrpL*; pBAD: *hrpL* were grown in hrp-repressing minimal media for a total of 22 hrs. The media was supplemented with 200 mM arabinose for 0, 1, 3, or 5, hours prior to RNA extraction. Qiagen's RNeasy extraction kit with Bacteria Protect was used and RNA was DNAse treated with Roche DNaseI in the presence of RNase Inhibitor.

Microarray and experimental design. A tiled array containing a total of 355,201 probes and covering 10X the draft genome of *P. carotovorum* WPP14 [32, 33] was designed by Jeremy Glasner. Probes are 40 to 60 bp long, with an average melting temperature of 79˚C and with an average spacing of 12 bp being alternated between the forward and reverse strands. The experiment was designed including 4 biological replications of bacterial RNA obtained from stem-bacterial interaction and 3 biological replications for the tuber-bacterial interaction.

cDNA synthesis and microarray hybridization. Bacterial RNA samples were submitted to the Expression Center at the Biotechnology Center at University of Wisconsin-Madison for cDNA synthesis, labeling, hybridization, washing, array scanning and data extraction. All the process was done following the Version 2.0 of NimbleGen user's guide for NimbleChip X1, which include the use of the Invitrogen Superscript Double-Stranded cDNA Synthesis Kit for cDNA synthesis and the 5' Cy3-labeled Random 50 Nanomers (9mer "Wobble") (TriLink Biotechnologies) for the labeling step. Labeled samples were hybridized to the *P. carotovorum* WPP14 1-plex NimbleExpress GeneChip array according to the NimbleGen Expression Protocol. Array scanning was done using the Axon 4000B scanner and the data were extracted using NimbleScan software.

Data analysis and statistical methods. Raw hybridization probe signals of all samples were normalized using quantile normalization as described by Bolstad, et al. [34] using NimbleScan software. Normalized raw signal values of each CDS were extracted from the data using Microsoft Access Software, then a median value of probes corresponding to each annotated feature were calculated. Partek Genomics Suite Software, version 6.4 (Partek Incorporated, St. Louis, Mo) was used to analyze the data by principal component assay (PCA), analysis of variance (ANOVA) and false discovery rate (FDR). PCA analysis was conducted to reduce dimensionality of the dataset and visually evaluate differences or similarities among biological replicates that allow them to

be grouped together. Differentially expressed genes (DEG) were calculated using ANOVA analysis with an additional FDR calculation to correct false positive rate. DEG were then listed using a cutoff value of 0.01. Axon 4000B scanner and the data were extracted using NimbleScan software.

Confocal Microscopy. Cultures were grown over night in rich media, washed twice in 10 mM $MgCl_2$ and re-suspended at an OD_{600} of 1.0 and infiltrated into leaves of 4-5 week old *A. thaliana* using a needle-less syringe. Potato leaves were not used because they were too thick to properly mount on microscope slides and visualize with confocal microscope. A Zeiss confocal microscope was used to visualize fluorescent bacteria in planta and images were captured and viewed with LSM imagebrowser software (Zeiss).

qRT-PCR done at UW. To validate the transcriptome profiling of the experiment, a quantitative real-time PCR (qRT-PCR) was carried out on selected candidate genes. Bacterial RNA was isolated as described above. Primers were designed based on the draft genome sequence of WPP14 using the BeaconDesigner software (Premier Biosoft International, Palo Alto, CA, U.S.A). Primers were designed in regions of little secondary template structure. Sequence primers are shown in Table 1. Primer efficiency was determined using dilution series of target DNA [36]. RNA samples were tested for residual DNA contamination using qRT-PCR and primers for targeting the *proC* gene in WPP14. RNA

samples that showed threshold cycles (Ct) values >30 cycles were considered to be sufficiently free of chromosomal DNA contamination.

cDNA synthesis was performed using iScript cDNA synthesis kit according to the manufacturer's instructions (Bio-Rad Laboratories, Hercules, CA, U.S.A.). Briefly, cDNA synthesis was performed with 500 ng of total RNA in 15 µl of DEPC-water, 4 µl of 5X iScript reaction mix that contains a blend of oligo dT and random hexamer primers and 1 µl of iScript reverse transcriptase. The reaction conditions were performed at 25 °C for 5 min., 42 °C for 30 min., and 85 °C for 5 min. Three independent biological replicates samples per experimental condition were used for the cDNA synthesis. Two iScript reactions were performed for each RNA sample and for each cDNA sample, two reactions were run in two wells of 25µl reaction. Expression of mRNA was determined by qRT-PCR using MyiQ detection system (Bio-Rad Laboratories). PCR conditions were 95°C for 3 min; 40 cycles of 95°C for 10 s and 500°C for 45 s; and 1 cycle of 95°C for 1 min and 55°C for 1 min; followed by a dissociation curve with 80 cycles of 55°C for 10 s with a 0.5°C increase per cycle. Primer-dimers and single presence of a product per reaction were evaluated using the dissociation (Melt) curve analysis built in the MyiQ detection system software.

For absolute quantification, target gene abundance was determined using standard curve estimation in each plate. This method allows estimation of the actual quantity of the mRNA in the sample avoiding differences in primer efficiency and product fluorescence. The mean of the starting quantity (SQ) ratios for each target gene were internally normalized using the absolute expression

mean values of gene *ffh*, which abundance also showed to be stable under different experimental conditions according to the Excel-based software program BestKeeper [36, 37]. In order to determine significant high or low expressed genes relative to the expression of the gene *hrpL*, a pair-wise comparison was performed to calculate *p*-values using the unpaired two-tailed t-test with 95% confidence interval using Prism 5.0a software (GraphPad Software, Inc.).

For relative expression analysis, target gene abundance was internally normalized to four reference genes (*dspE, pelB, hrcC and ffh*) which were previously checked for mRNA stability under different experimental conditions using the BestKeeper program. Target transcripts amounts were reported as relative expression ratio (RER) of target transcript in stems relative to tubers. RER was estimated as the difference between Ct values which was calculated using the equation 2–ΔΔCt as previously modified [36, 38]. Statistical analysis of RER values of selected genes were conducted using the unpaired two-tailed t-test with 95% confidence interval using Prism 5.0a software (GraphPad Software, Inc.).

Table 1. qRT-PCR primer sequences and efficiency

Gene name	Primer name	Primer sequence (5'→3', forward to reverse)	Efficiency (%)A
Ffh	P0	TGGAAACATTGGCAGAGC	…
		GACTAACAAGACATCGTAGAAC	105
BudB	P0	TTGAATCTGCTGATGAAC	…
		CAATGGTTATCGGAATAATC	99.5

NarI	P0	TTGACATCTATCCTTACCT	…
		CCTTTCTTATCCAGCATT	93.2
Evr	P0	GCTTCTAATAAACAAATCA	…
		CGTTCTTCTTCTAATAGTA	98.2
pelB	P0	CTCCGTAACAACAACATT	…
		GTACTCTTCCAGTCATCT	102.3
fliC	P0	CGAATCTACCATTACTAACCT	…
		AGTCAGCGTCTTCAATAC	98.5
dspE	P0	GTCCTATACCAACCTCAG	…
		GCAACGAAGAGAACAAAT	98.8
hrpN	P0	CAGGAGTTGAACAACATTAG	…
		CCATCTTACGGTCTTCTT	88.8
hrpL	P0	AATGACCTATCTGGAAGT	…
		TTGAAATAATTGCGAACC	95.3
hrpW	P0	AGGTAACTGGAACTTGAA	…
		TTCGTCAGGTTGATGTTA	95.6
hrcC	P0	GCCTATTCTGCCGAACAA	…
		GCACCAAATCCACACCAT	99.9

A: Primer efficiencies were calculated from dilution curve using the MyiQ Clycer software

qRT-PCR done at UNC. RNA was reverse transcribed into cDNA using the Ambion RETROscript kit. Briefly, 2ug of RNA was primed with random

decamers and reverse transcribed with the MMLV-RT. cDNA was then diluted and relative quantities of specific transcripts were determined using SYBR Green RT-PCR Reagents. Fluorescence of double-stranded DNA is measured by a DNA engine, Opticon 2 Continuous Fluorescence Detector (MJ Research) and values are analyzed using Opticon Monitor 3 software (MJ Research). Relative gene expression is determined by normalizing to a *P. carotovorum* house-keeping gene, *ffh,* which encodes a signal recognition particle protein, and was identified as an optimally stable control gene for qRT PCR analysis by Takle, Toth, and Brurberg [39].

RESULTS AND DISCUSSION:

Virulence assays indicate that the HrpL regulon is required for virulence. In order to determine the contribution of the T3SS on virulence, we adapted a leaf virulence assay and compared wild-type growth in leaves to that of a *hrpL* mutant. Briefly, 4-5 week old *A. thaliana* and Yukon gold potato leaves were infiltrated with a dilute culture (10^4 CFU/mL) of *P. carotovorum* by a needle-less syringe. Figure 1a and 1b show that in three days after infiltration, in both plants, the *hrpL* mutant was severely attenuated in virulence with 3 logs less growth compared to wild type. We also saw a reduced incidence of maceration symptoms in the leaves of both hosts (data not shown). Here we've established *A. thaliana* as a model host, in that it mimic the disease and growth patterns of *P. carotovorum* in the native host, potato. Thus, *A. thaliana* is a relevant host and is

used for other analysis in this chapter and we suggest it as a useful tool for future work on *P. carotovorum* pathogenesis. Given this pathology data on a native and model host, we found that the HrpL-regulon, and by association, T3SS-related virulence factors, are required for full disease and growth on host leaves.

Figure 1

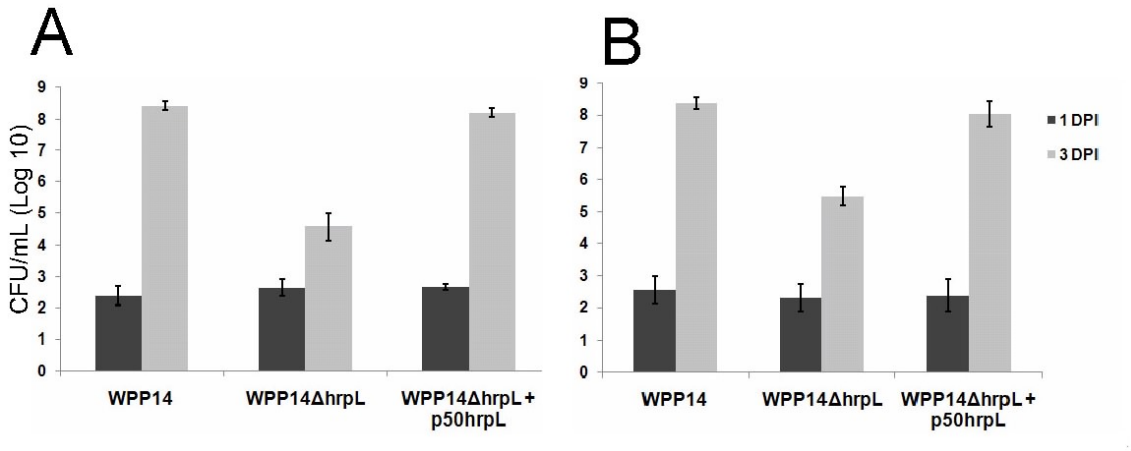


Figure 1. ***hrpL* mutant is defective in growth on leaves**. Potato (A) and *A. thaliana* (B) hand infiltration of *Pcc* WPP14, the *hrpL* operon, or the *vgu* operon mutant carrying pC50:*hrpL* with 10^4 CFU/mL initial inoculum, into leaves of 4-5 week old plants. Experiment has four internal replicates and was repeated three times; error bars indicate two times standard error.

Interestingly, when we infected potato tubers with wild-type and mutant and measure the area of macerated tissue, we saw no difference between a wild type *P. carotovorum* and a *hrpL* mutant [40], suggesting that the role of the HrpL-

regulon in virulence is specific to the given infection route and that a different virulence strategy is required for tuber infection.

Promoter-trap screen to identify HrpL regulated genes. In order to identify additional T3SS effectors and novel T3SS-associated virulence factors, we adapted a FACS-based promoter trap screen developed by Chang et al. 2005 [28] to identify T3S effectors in *Pseudomonas syringae*. Previous studies of the T3SS in *P. carotovorum* had identified the alternative sigma factor, HrpL, as the master regulator of the T3SS and the known effector, DspE [21]. Additionally, in *P. syringae*, a homologous HrpL is known to regulate the expression of all known effectors. Therefore, we exploited HrpL as a likely regulator of unidentified T3S effectors and associated virulence factors. We cloned *hrpL* from *P. carotovorum* WPP14 into a pBAD vector, with *hrpL* expression under the regulation of the *araBAD* promoter, which is inducible with arabinose. We then constructed a genome library of WPP14 in a differential fluorescent induction vector (DFI), where the library fragments were cloned downstream of a promoter-less *gfp* gene [28]. The library fragments ranged from 1-4kb in size using either *Tsp509I* (an isoschizomer of *EcoR1*) or a set of blunt enzymes (*AluI*, *BstUI*, *HaeIII*, and *RsaI*). *dspE* was also cloned into the DFI vector to serve as a positive control for induction, and peak fluorescence after *hrpL* induction. *dspE* and the library clones in the DFI vector were mobilized by tri-parental mating into WPP14, which already carried *hrpL* on the arabinose-inducible pBAD plasmid. Cultures were

sorted by FACS with 4 tandem sorts to enrich for *hrpL*-dependent fluorescing cells.

Approximately 2600 clones were captured and their library fragments were sequenced and aligned to the genome. Sequences aligned to 37 regions (<4kb in length) of the genome. 17% of the sequences aligned to an arabinose operon responsive to the induction conditions and were therefore disregarded. Randomly selected clones representing each region were independently grown with and without induction and re-screened by FACS to verify HrpL-dependent fluorescence. We found a 20% false-positive rate (determined by the number of sequences that aligned to false genome regions (not HrpL-inducible) divided by the total number of sequences) and verified 17 representative genome regions as HrpL-inducible. Genome region information and the direction of the library fragment in the DFI vector allowed us to identify the full length ORF that was indicated as HrpL-regulated. Genes identified as HrpL-dependent by FACS re-screening were also verified by using qRT-PCR to determine HrpL-dependent expression of the native gene. This identified additional false-positives for a total of 25% false-positives coming out of the FACS screen. False-positives identified this way appeared to have HrpL-induction as an artifact of cloning and fusing to *gfp* in the DFI vector.

The 6 genome regions identified and verified (representing 58% of the sequenced library clones) all aligned to known HrpL-dependent genes and operons within the cluster of genes that encode the T3S apparatus (Table 2). The region that was represented the most, with 784 clones identified, was the 12

member operon within the T3SS cluster. Known *hrp* cluster members and predicted HrpL binding sites were based on the Lehtimaki *et al.* 2003 published description of the T3SS cluster [41]. *dspE*, the one known effector gene previously identified, was represented in the screen 329 times. *hrpA*, encoding a T3SS-associated pilin, and *hrpF*, a putative T3SS translocon, were represented the least with only 6 and 4 representative clones, respectively. They appear to be lowly and constitutively expressed and therefore not optimally found in this type of screen. Only one known HprL-regulated gene was not identified in the screen, *hrpN*. We cloned *hrpN* into the DFI vector to assess its fluorescent profile with and without *hrpL* induction from pBAD:*hrpL* using FACS. We found that *hrpN* expression was inducible via HrpL, however its basal expression level was high (higher than 30% of the fluorescence of the library) and therefore was lost during the screen. Additionally, HrpN is an exceptional harpin known to be regulated by multiple transcription factors in *P. carotovorum* and close relative *E. chrysanthemi*, including regulators KdgR and PecS [42-44].

Thus, although the HrpL-promoter trap screen did not identify any new HrpL-regulated genes that could be candidates for T3S effectors or associated virulence factors, it did identify known HrpL-regulated genes, validating the utility and concept of the screen in *Pectobacterium* species.

Table 2:

Gene	Function	Number of clones	Minimum number of independent inserts
hrpQ/J **ADT-0002852**	**Type III secretion operon**	**784**	**74**
hrpA **ADT-0001735**	**Type III secretion operon**	**5**	**2**
hrpF **ADT-0001736**	**Type III secretion operon**	**6**	**2**
hrpW **ADT-0002859**	**Type III secretion accessory protein**	**149**	**16**
hrpK **ADT-0002735**	**Type III secretion accessory protein**	**196**	**17**
dspE **ADT-0003862**	**Type III effector**	**329**	**19**

Table 2: HrpL regulated genes found in the DFI screen. T3SS accessory genes are involved in apparatus assembly. When a *hrp* cluster operon was identified, the first gene of the operon is listed.

Transcriptional profiling identifies the HrpL regulon. To further explore the HrpL regulon, we employed a tiled NimbleGen array to identify up- and down-regulated members of the HrpL regulon. We compared the transcription profiles

of a *hrpL* mutant carrying an empty pBAD vector and a *hrpL* mutant strain over-expressing *hrpL* from the pBAD promoter. Based on FACS data from the previously discussed screen, we found that the activity of HrpL-responsive promoters peaked after 4 hours of induced HrpL expression from pBAD:hrpL (Figure 2). Therefore, we profiled transcription of both strains after 0, 1, 3 and 5 hours of arabinose induction (with a total growth time of 22 hours for all induction time points). Array data was analyzed using EBarrays package in R statistical environment [45, 46]. EBarrays provides three different distribution models with posterior probability calculated to account for false-discovery rate. Using all three models from this package, we found that 100 genes were differentially expressed between the two strains throughout the time course. Thirty seven genes were identified in all three time points and they are shown in Figure 3. Twenty nine of the 37 genes that were identified in all the time points are genes known to be involved in T3SS, most of which are found within

Figure 2.

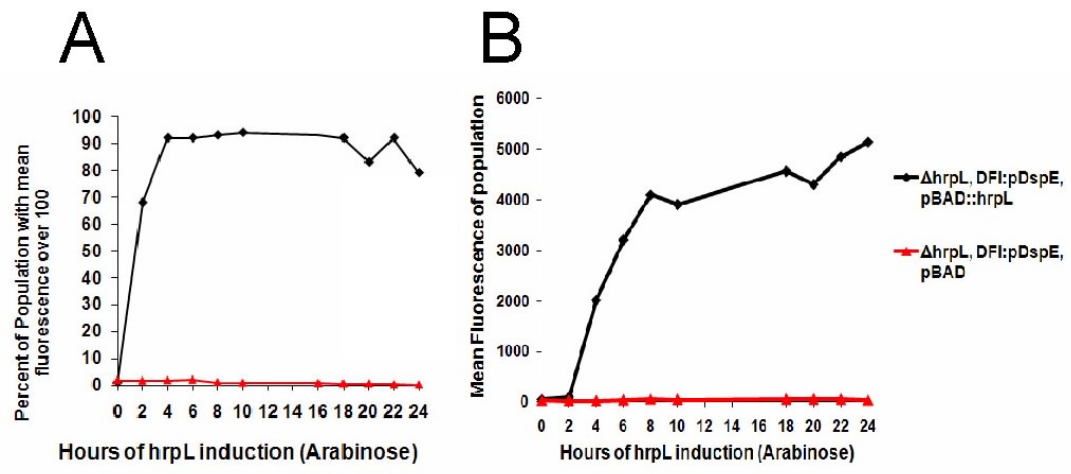


Figure 2: Graphs showing the mean fluorescence of *P. carotovorum* when expressing GFP from the promoter of *dspE* in the presence of 200 mM arabinose and either pBAD or pBAD:hrpL. Figure 2A shows the percentage of the population that has fluorescence over 100, a point determined to be definitively above background fluorescence. Figure 2B shows the mean fluorescence of the whole population. Time points indicate the duration of arabinose induction. All cultures were grown in minimal media for 18 hours.

---

the T3SS gene cluster. HrpN was also identified in this cluster of genes, however, it has higher basal expression level in the absence of HrpL over-expression. The remaining 8 genes identified as differentially expressed in all three time points were previously not known to be regulated by HrpL or to be associated with the T3SS.

Two of the 8 genes were down-regulated by HrpL induction. These two genes, a putative hexuronate transporter and a putative glycosyl hydrolase, were shown by Van Gijsegem *et al.* [47] to be involved in virulence on African violets and up-regulated by a *lacI*-like repressor, designated LfaR in *Erwinia chrysanthemi*. LfaR, however, was not identified as differentially expressed between the hrpL mutant and the hrpL over-expression strain.

The remaining 6 novel genes found to be up-regulated by HrpL include a protein with unknown function, designated *ytfK* in a wide-variety of other Enterobacteria, and a putative oxidoreductase. HrpL up-regulates a 59 aa putative exported protein, which is only found in *Pectobacterium* species and

Figure 3.

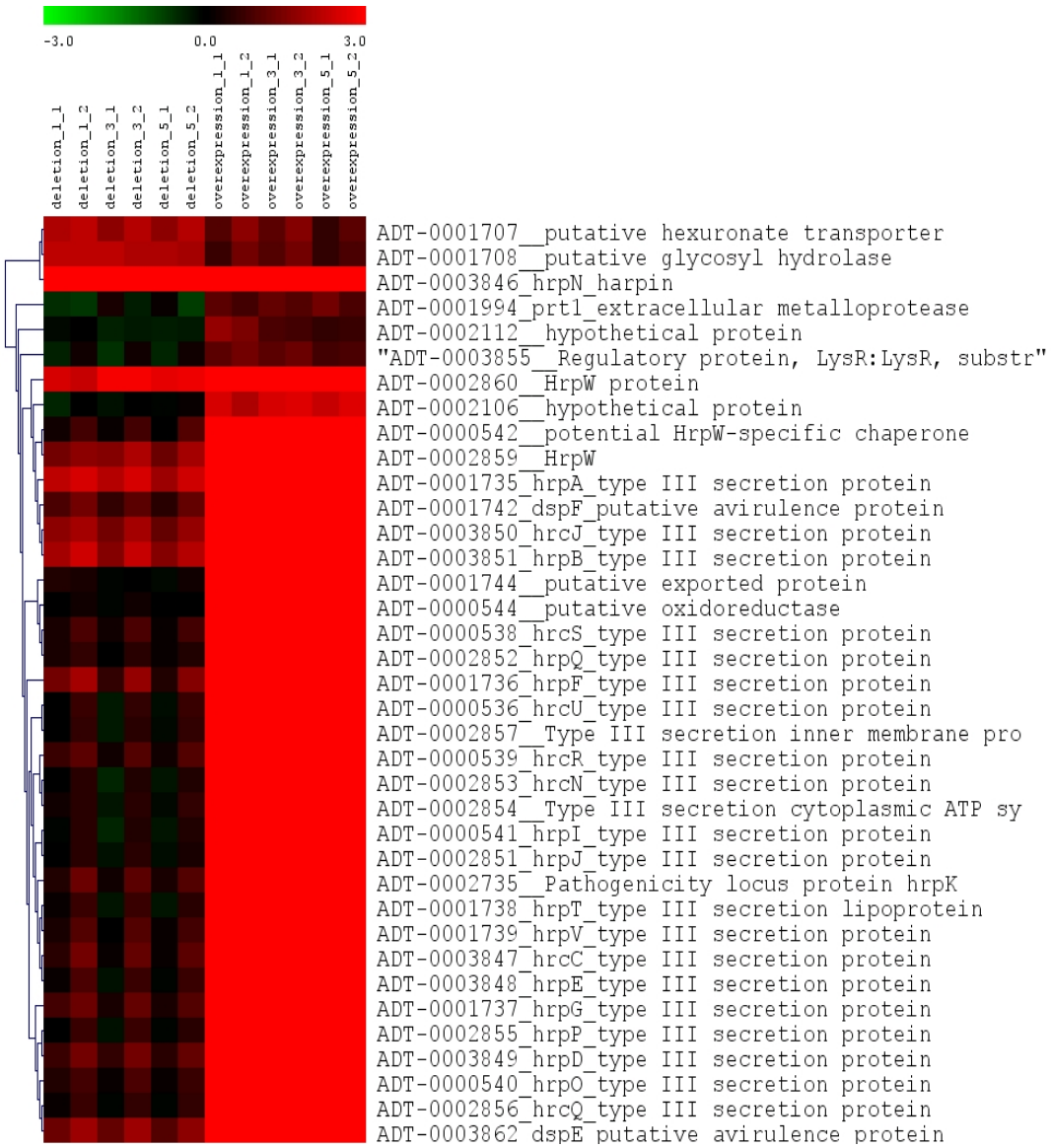


Figure 3: Thirty-seven genes identified in the HrpL regulon. A heat map made using MultiExperiment Viewer indicating the relative expression of genes found to be significantly differentially expressed between the two strains in all of the time points. Green indicates relatively low expression, black indicates neutral expression, and red indicates relatively high expression. The gene designations

from their ASAP annotations (introduced in Chapter 2) are shown on the right. The genes ordered in hierarchal clustering based on their expression patterns

---

Dickeya zeae. A poorly-understood metalloprotease, *prt1*, was also identified as up- regulated by HrpL. Prt1 is known to be involved in disease and expressed *in planta*, but its exact function remains unclear. Lastly, two regulators were identified: One is a *lysR*-like regulator located within the T3SS cluster of genes, suggesting it may have a secondary role of regulation for associated T3SS genes; the other is a hypothetical protein, found to contain GGDEF and GAF domains, which are each implicated in c-di-GMP signaling. Second messengers such as cAMP and c-di-GMP have been well established in other organisms as critical regulators of a phenotypic switch between motility and virulence. There has previously been no work on small molecule signaling in *P. carotovorum*, but this finding suggests it may use c-di-GMP as a second messenger during pathogenesis.

Table 3

Genes Differentially Expressed 1 Hr After Induction	FOLD CHANGE	
ADT-0000817_galK_galactokinase	2	down
ADT-0000589_trpE_anthranilate synthase component	2	down
ADT-0000818_galT_galactose-1-phosphate uridylyltra	2	down
ADT-0000590_trpG_anthranilate synthase component	2	down
ADT-0003084__Phospho-2-dehydro-3-deoxyheptonate	2	down
ADT-0003085__Phospho-2-dehydro-3-deoxyheptonate	2	down
ADT-0003927__Outer membrane protein F precursor	2	down

ADT-0000591_trpD_anthranilate phosphoribosyltransferase	2	down
ADT-0002737__Beta-xylosidase	2	down
ADT-0002884__Putative sporulation protein	3	down
ADT-0002885__Putative sporulation protein	4	down
ADT-0001287__succinate-semialdehyde dehydrogenase	2	up
ADT-0002736__Xyloside transporter XynT	2	up
ADT-0003752__Transcriptional regulator, AraC family	3	up
ADT-0003861__HecB precursor	5	up
ADT-0002861__Probable transcriptional regulator	6	up

Genes Differentially Expressed 3 Hr After Induction	FOLD CHANGE	
ADT-0002358__methyl-accepting chemotaxis protein	2	down
ADT-0000886__hypothetical protein	2	down
ADT-0004458_tRNA-Ile_tRNA-Ile(GAU)	2	down
ADT-0002429__hypothetical protein	2	down
ADT-0000027__methyl-accepting chemotaxis protein	2	down
ADT-0000048_metE_5-methyltetrahydropteroyltrigluta	3	down
ADT-0001893_fruK_1-phosphofructokinase	4	down
ADT-0003384__putative short chain dehydrogenase	2	up
ADT-0000446__hypothetical protein	2	up
ADT-0004037_ibpB_heat shock protein B	2	up

Genes Differentially Expressed 5 Hr After Induction	FOLD CHANGE	
ADT-0004008_prtW_metalloprotease	3	down
ADT-0000886__hypothetical protein	3	down
ADT-0002429__hypothetical protein	3	down
ADT-0000027__methyl-accepting chemotaxis protein	2	down

ADT-0000884__hypothetical protein	2	down
ADT-0000430_fliC_flagellin	2	down
ADT-0001964__hypothetical protein	2	down
ADT-0000885__hypothetical protein	2	down
ADT-0002358__methyl-accepting chemotaxis protein	2	down
ADT-0004237__hypothetical protein; Hypothetical secreted protein	2	down
ADT-0000879__putative membrane protein	2	down
ADT-0000881__putative lipoprotein	2	down
ADT-0000883__hypothetical protein	2	down
ADT-0003621__Methyl-accepting chemotaxis protein I	2	down
ADT-0002772_cheD_methyl-accepting chemotaxis protein	2	down
ADT-0002654__probable exported protein YPO2987	2	down
ADT-0000882__hypothetical protein	2	down
ADT-0002777_flgK_flagellar hook-associated protein	2	down
ADT-0000435_fliZ_putative alternative sigma factor	2	down
ADT-0000428_flgL_flagellar hook-associated protein	2	down
ADT-0003787__Flagellar motor switch protein fliG	2	down
ADT-0001641_fliF_flagellar M-ring protein	2	down
ADT-0002771__Methyl-accepting chemotaxis protein	2	down
ADT-0003012__Methyl-accepting chemotaxis protein	2	down
ADT-0000878__putative chaperone	2	down
ADT-0001767__methyl-accepting chemotaxis protein	2	down
ADT-0002861__Probable transcriptional regulator	4	up
ADT-0002065__putative exported protein	3	up
ADT-0001823__hypothetical protein	3	up
ADT-0000644_pepT_peptidase T	3	up
ADT-0000772__hypothetical protein	2	up
ADT-0003884_ompW_outer membrane protein W	2	up

ADT-0004037_ibpB_heat shock protein B	2	up
ADT-0004254_nirB_nitrite reductase [NAD(P)H] large	2	up
ADT-0000446__hypothetical protein	2	up
ADT-0004038_ibpA_heat shock protein A	2	up
ADT-0002217_nirD_nitrite reductase [NAD(P)H] small	2	up
ADT-0002423__hypothetical protein	2	up
ADT-0003804__hypothetical protein	2	up
ADT-0000344__putative membrane protein	2	up
ADT-0000771__putative haloacid dehalogenase-like	2	up
ADT-0003572__putative exported protein	2	up
ADT-0003263_aegA_anaerobically expressed oxidoreductase	2	up
ADT-0001521__hypothetical protein	2	up

Table 3: Genes found to be in the HrpL stimulon: Gene identification numbers were determined by the annotation described in Chapter 2. Highlighting indicates genes that were found in two time points.

Genes found to be differentially expressed in only one or two of the time points are shown in Table 3. Interestingly, in the later time points, there are multiple methyl-accepting chemotaxis proteins and flagellar genes down regulated. It has been previously reported in other bacteria that motility was down-regulated while the T3SS is expressed [48-52]. This, however, is the first instance documented in *P. carotovorum* which indicates the master-regulator of T3SS, HrpL, is involved in down-regulating motility.

Together, this data indicates that the HrpL-regulon of *P. carotovorum* is small relative to those of *P. syringae*, the plant-pathogen with the

phylogenetically closest T3SS. However, multiple known virulence regulators were identified here as downstream of HrpL activity, suggesting a broader role for HrpL in virulence regulation and an integration of the T3SS into the complex virulence regulation network of *P. carotovorum*.

Confocal Microscopy to confirm HrpL activity in vivo. To complement our pathology data of the HrpL regulon's role *in planta*, we used confocal microscopy to visualize the activity of a HrpL-responsive reporter during leaf infection. Here we used a *dspE* promoter fusion to *gfp* (in our DFI vector) from our promoter – trap screen. This construct was carried by wild-type *P. carotovorum*, which was then infiltrated into *A. thaliana* leaves. Here, *A. thaliana* leaves were used due to their relative thinness compared to potato leaves, which made it easier to visualize fluorescent bacteria with confocal microscopy. In Figure 4, fluorescence is visualized 7 hours after infiltration, indicating native HrpL induction and activity and T3SS expression *in vivo*. Importantly, also in Figure 4, when our $dspE_{pro}$-gfp fusion was carried in a *hrpL* mutant, we do not see this fluorescence at any of the sampled time points, indicating that the fluorescence we saw before was due to HrpL activity. This again indicates that the T3SS is expressed *in planta* and is expressed early in infection. Later time-points using confocal microscopy were not taken due to tissue maceration and for fear that the stability and long half-life of GFP would skew interpretations of the timing of expression.

Quantitative Real-Time PCR indicates that HrpL and the T3SS are expressed early in infection. To understand specific stages of leaf infection, we used quantitative RT-PCR to look at expression of key virulence factors and regulators during a time course spanning early infection. Bacterial RNA was extracted from infected potato leaves at various time points and relative

Figure 4:

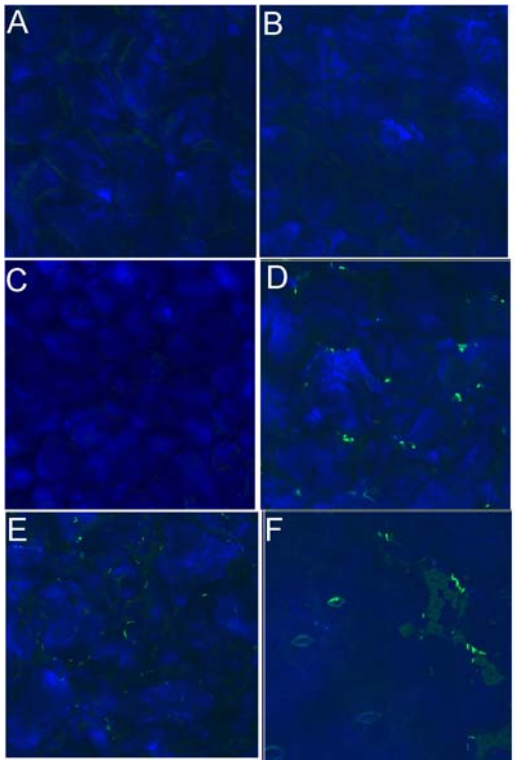


Figure 4. Confocal microscopy images of *A. thaliana* mesophyl cells 7 hours after infection with *P. carotovorum.* A. mesophyl cells with no *P. carotovorum*. B. mesophyl cells infected with WPP14. C. Mesophyl cells infected with WPP14Δh*rpL*, DFI:$dspE_{pro}$GFP as a negative control. D. Mesophyl cells infected with 200 mM Arabinose and WPP14, DFI:$dspE_{pro}$GFP and pBAD:*hrpL* as a positive control. E and F. Mesophyl cells infected with WPP14, DFI:$dspE_{pro}$GFP.

expression of *hrpL*, the effector *dspE*, harpins *hrpN* and *hrpW, pelB* and *fliC* were selected to generally observe the timing of expression of the T3SS, T2SS, and motility during early leaf infections. *hrpN* and *hrpW* encode harpin proteins that are secreted, but not translocated, by the T3SS and known to elicit defense responses in *A. thaliana*. HrpN is regulated by HrpL, but also by repressors that respond to plant signals (KdgR and PecT). *pelB* encodes a pectate lyase, secreted via the T2SS. *fliC* encodes the flagellin subunit. Figure 5 shows qRT-PCR data of gene expression 7 hours after infection of potato leaves. *hrpL* and genes downstream of *hrpL, dspE* and *hrpW* are up-regulated, while *hrpN,* which is regulated by multiple transcription factors, is not seen to be up-regulated. It is surprising that this harpin is not co-expressed with the T3SS here since the locus is within the hrp cluster and is regulated by HrpL. However, transcriptional data from tubers and stems also indicate that HrpN is expressed independently from the T3SS, which may be due to the action of *hrpN* repressors [42, 43]. Similar to *hrpN*, *pelB and fliC* are also lowly expressed during early leaf infection. This data suggests that while the T3SS is expressed, *P. carotovorum* are non-motile and not secreting macerating enzymes via a T2SS. This observation is in accordance with the down-regulaton of flagellar genes in later time points of the DNA array (described above) and with transcriptional profiling of *P. carotovorum* done in stem and tubers, where we see motility and T2SS genes relatively highly expressed and T3SS genes relatively lowly expressed. Before the data in this chapter is submitted for publication, this experiment will be repeated with a 0 time

control to better assess relative expression levels. However, from preliminary data where 0 time points were taken, we know that *hrpL* and *dspE* are not expressed in the minimal media used prior to leaf infiltration.

Figure 5

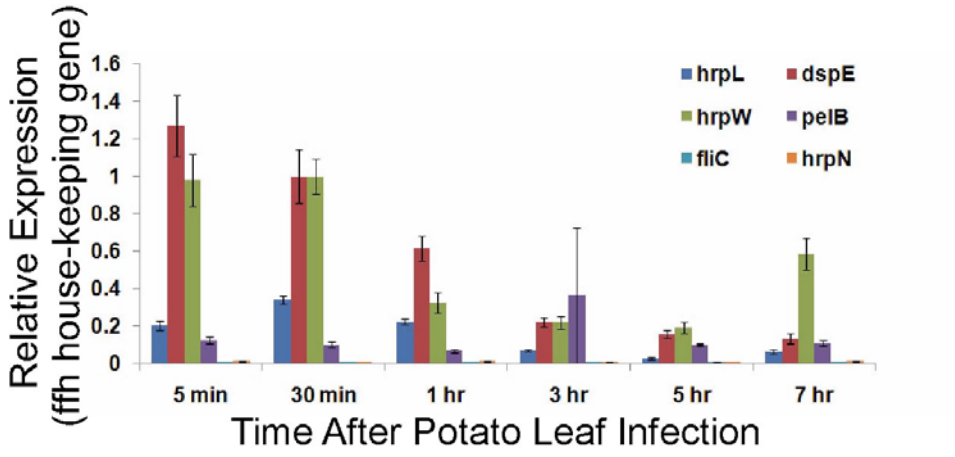


Figure 5: Relative expression of key virulence regulators and factors during early leaf infection. The expression of the T3SS-regulator, *hrpL* (dark blue) the T3SS helper genes, *hrpW* (green) and *hrpN* (orange), the T3SS effector-encoding gene, *dspE* (red), and the gene encoding the flagellin subunit, *fliC* (light blue), and the pectate lyase-encoding gene, *pelB* (purple). Expression is relative to the house-keeping gene, *ffh.* Error bars represent standard error. This experiment needs to be repeated.

Microarray data of transcription profiles during tuber and stem infection. My collaborators Amy Charkowski and María del Pilar Márquez used transcriptional profiling of WPP14 during tuber and stem infections to identify virulence factors important for stem and tuber infection routes. The microarray

data analysis and verifying qRT-PCR are still being completed. Preliminary results indicate that the transcriptional profiles of WPP14 during stem and tuber infection were largely the same and genes relatively highly expressed in both stem and tuber infections principally included those involved in motility and secretion of plant cell wall degrading enzymes, including the T2SS apparatus encoding genes. HrpL and the T3SS-associated genes were weakly expressed in both stem and tuber infections. A set of 139 genes were found to be differentially expressed between the two infection routes: 74 were specifically highly expressed in stem, and 65 in tuber infections. Genes specifically induced during the WPP14-stem interaction include a nitrate transport and metabolism operon, an osmotically inducible lipoprotein, the global regulator *rpoS*, a negative regulator of multidrug resistance, and a phosphate transport system. Genes specifically induced during the WPP14-tuber interaction include a glycerol-3-phosphate transport system, a putative type IV pilus protein, a metallopeptidase, and a two-component signal transduction system involved in citrate catabolism. Although more analysis and discussion is required, the data generally indicates that the T3SS is not expressed during these infection conditions, macerating enzyme production and motility are necessary, and that each interaction also produces tissue-specific virulence and fitness-related expression patterns.

CONCLUSION:

In this chapter, we explored the transcriptional profiles of *P. carotovorum* and the role of T3SS during infection. We found that the T3SS and associated virulence

factors are required for virulence on the leaves of *A. thaliana* and potato. In accordance with this finding, we used confocal microscopy to visualize *in planta* expression of the T3SS during leaf infections and found that levels of GFP could be seen after 7 hours of infection. We used a promter-trap screen and DNA microarray to determine that HrpL, the master regulator for the T3SS, regulates only 8 genes in addition to known T3SS hrp cluster genes. Further transcriptional analysis indicates that the T3SS is expressed very early upon infection in leaves. These results suggest a tissue- and timing-specific role for the T3SS during pathogenesis and that the T3SS of *P. carotovorum* has a relatively small HrpL regulon and has been incorporated into the pathogenesis strategy differently than in *P. syringae*, which has a similar system.

LITERATURE CITED:

1. McCarter-Zorner, N.J., Franc, G. D., Harrison, M. D., Michaud, J. E., Quinn, C. E., Sells, I. A., and Graham, D. C., *Soft rot Erwinia in surface and underground waters in Southern Scotland and in Colorado, United States.* J. Appl. Bacteriol., 1984. **57**: p. 95-105.

2. McCarter-Zorner, N.J., Harrison, M. D., Franc, G. D., Quinn, C. E., Sells, I. A., and Graham, D. C., *Soft rot Erwinia bacteria in the rhizosphere of weeds and crop plants in Colorado, United States and Scotland.* J. Appl. Bacteriol. , 1985. **59**: p. 357-368.

3. Harrison, M., Quinn, C., Ann Sells, I., and Graham, D., *Waste potato dumps as sources of insects contaminated with soft rot coliform bacteria in relation to re-contamination of pathogen-free potato stocks.* Potato Res., 1977. **20**(1): p. 37-52.

4. Molina, J.M., Harrison, D., and Brewer, J. W., *Transmission of Erwinia carotovora var. atroseptica by Drosophila melanogaster Meig. I. Acquisition and transmission of the bacterium.* Am. Potato J., 1974. **51**: p. 245-250.

5. Allen, C., S. Reverchon, and J. Robert-Baudouy, *Nucleotide sequence of the Erwinia chrysanthemi gene encoding 2-keto-3-deoxygluconate permease.* Gene, 1989. **83**(2): p. 233-41.

6. Matsumoto, H., et al., *Comparative study of regulatory mechanisms for pectinase production by Erwinia carotovora subsp. carotovora and Erwinia chrysanthemi.* Mol. Plant Microbe Interact., 2003. **16**(3): p. 226-37.

7. Lagaert, S., T. Belien, and G. Volckaert, *Plant cell walls: Protecting the barrier from degradation by microbial enzymes.* Semin. Cell Dev. Biol., 2009. **20**(9): p. 1064-73.

8. Rojas, C.M., et al., *HecA, a member of a class of adhesins produced by diverse pathogenic bacteria, contributes to the attachment, aggregation, epidermal cell killing, and virulence phenotypes of Erwinia chrysanthemi EC16 on Nicotiana clevelandii seedlings.* Proc Natl Acad Sci U S A, 2002. **99**(20): p. 13142-7.

9. Lopez-Solanilla, E., F. Garcia-Olmedo, and P. Rodriguez-Palenzuela, *Inactivation of the sapA to sapF locus of Erwinia chrysanthemi reveals common features in plant and animal bacterial pathogenesis.* Plant Cell, 1998. **10**(6): p. 917-24.

10. Marits, R., et al., *Regulation of the expression of prtW::gusA fusions in Erwinia carotovora subsp. carotovora.* Microbiology, 2002. **148**(Pt 3): p. 835-42.

11. Toth, I.K. and P.R. Birch, *Rotting softly and stealthily.* Curr. Opin. Plant Biol., 2005. **8**(4): p. 424-9.

12. Valls, M., S. Genin, and C. Boucher, *Integrated regulation of the type III secretion system and other virulence determinants in Ralstonia solanacearum.* PLoS Pathog, 2006. **2**(8): p. e82.

13. Francis, M.S., H. Wolf-Watz, and A. Forsberg, *Regulation of type III secretion systems.* Curr Opin Microbiol, 2002. **5**(2): p. 166-72.

14. Cunnac, S., M. Lindeberg, and A. Collmer, *Pseudomonas syringae type III secretion system effectors: repertoires in search of functions.* Curr Opin Microbiol, 2009. **12**(1): p. 53-60.

15. Nomura, K., M. Melotto, and S.Y. He, *Suppression of host defense in compatible plant-Pseudomonas syringae interactions.* Curr Opin Plant Biol, 2005. **8**(4): p. 361-8.

16. Grant, S.R., et al., *Subterfuge and manipulation: type III effector proteins of phytopathogenic bacteria.* Annu Rev Microbiol, 2006. **60**: p. 425-49.

17. Belkhadir, Y., R. Subramaniam, and J.L. Dangl, *Plant disease resistance protein signaling: NBS-LRR proteins and their partners.* Curr Opin Plant Biol, 2004. **7**(4): p. 391-9.

18. Alfano, J.R. and A. Collmer, *The type III (Hrp) secretion pathway of plant pathogenic bacteria: trafficking harpins, Avr proteins, and death.* J Bacteriol, 1997. **179**(18): p. 5655-62.

19. Chisholm, S.T., et al., *Host-microbe interactions: shaping the evolution of the plant immune response.* Cell, 2006. **124**(4): p. 803-14.

20. Tampakaki, A.P., et al., *Conserved features of type III secretion.* Cell Microbiol, 2004. **6**(9): p. 805-16.

21. Chatterjee, A., Y. Cui, and A.K. Chatterjee, *Regulation of Erwinia carotovora hrpL$_{Ecc}$ (sigma-L$_{Ecc}$), which encodes an extracytoplasmic function subfamily of sigma factor required for expression of the HRP regulon.* Mol. Plant Microbe Interact., 2002. **15**(9): p. 971-80.

22. Mole, B.M., et al., *Global virulence regulation networks in phytopathogenic bacteria.* Trends Microbiol., 2007. **15**(8): p. 363-71.

23. He, S.Y., K. Nomura, and T.S. Whittam, *Type III protein secretion mechanism in mammalian and plant pathogens.* Biochim Biophys Acta, 2004. **1694**(1-3): p. 181-206.

24. Tang, X., Y. Xiao, and J.M. Zhou, *Regulation of the type III secretion system in phytopathogenic bacteria.* Mol Plant Microbe Interact, 2006. **19**(11): p. 1159-66.

25. Ferreira, A.O., et al., *Whole-genome expression profiling defines the HrpL regulon of Pseudomonas syringae pv. tomato DC3000, allows de novo reconstruction of the Hrp cis clement, and identifies novel coregulated genes.* Mol Plant Microbe Interact, 2006. **19**(11): p. 1167-79.

26. Toth, I.K. and P.R. Birch, *Rotting softly and stealthily.* Curr Opin Plant Biol, 2005. **8**(4): p. 424-9.

27. Sambrook, J. and D.W. Russell, *Molecular cloning : a laboratory manual*. 3rd ed. 2001, Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

28. Chang, J.H., et al., *A high-throughput, near-saturating screen for type III effector genes from Pseudomonas syringae.* Proc. Natl. Acad. Sci. U.S.A., 2005. **102**(7): p. 2549-54.

29. Datsenko, K.A. and B.L. Wanner, *One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products.* Proc. Natl. Acad. Sci. U.S.A., 2000. **97**(12): p. 6640-5.

30. Yap, M.N., J.D. Barak, and A.O. Charkowski, *Genomic diversity of Erwinia carotovora subsp. carotovora and its correlation with virulence.* Appl Environ Microbiol, 2004. **70**(5): p. 3013-23.

31. Glasner, J.D., et al., *Niche-specificity and the variable fraction of the Pectobacterium pan-genome.* Mol. Plant Microbe Interact., 2008. **21**(12): p. 1549-60.

32. Jahn, C.E., D.K. Willis, and A.O. Charkowski, *The flagellar sigma factor fliA is required for Dickeya dadantii virulence.* Mol Plant Microbe Interact, 2008. **21**(11): p. 1431-42.

33. Glasner, J.D., et al., *ASAP, a systematic annotation package for community analysis of genomes.* Nucleic Acids Res, 2003. **31**(1): p. 147-51.

34. Bolstad, B.M., et al., *A comparison of normalization methods for high density oligonucleotide array data based on variance and bias.* Bioinformatics, 2003. **19**(2): p. 185-93.

35. Lin-Chao, S. and H. Bremer, *Effect of the bacterial growth rate on replication control of plasmid pBR322 in Escherichia coli.* Mol Gen Genet, 1986. **203**(1): p. 143-9.

36. Pfaffl, M.W., et al., *Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations.* Biotechnol Lett, 2004. **26**(6): p. 509-15.

37. Vandesompele, J., et al., *Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes.* Genome Biol, 2002. **3**(7): p. RESEARCH0034.

38. Rotenberg, D., et al., *Methods for effective real-time RT-PCR analysis of virus-induced gene silencing.* J Virol Methods, 2006. **138**(1-2): p. 49-59.

39. Takle, G.W., I.K. Toth, and M.B. Brurberg, *Evaluation of reference genes for real-time RT-PCR expression studies in the plant pathogen Pectobacterium atrosepticum.* BMC Plant Biol., 2007. **7**: p. 50.

40. Kim, H.S., et al., *Phylogeny and virulence of naturally occurring type III secretion system-deficient Pectobacterium strains.* Appl Environ Microbiol, 2009. **75**(13): p. 4539-49.

41. Lehtimaki, S., et al., *Characterization of the hrp pathogenicity cluster of Erwinia carotovora subsp. carotovora: high basal level expression in a mutant is associated with reduced virulence.* Mol Genet Genomics, 2003. **270**(3): p. 263-72.

42. Nasser, W., et al., *PecS and PecT coregulate the synthesis of HrpN and pectate lyases, two virulence determinants in Erwinia chrysanthemi 3937.* Mol. Plant Microbe Interact., 2005. **18**(11): p. 1205-14.

43. Liu, Y., et al., *kdgREcc negatively regulates genes for pectinases, cellulase, protease, HarpinEcc, and a global RNA regulator in Erwinia carotovora subsp. carotovora.* J Bacteriol, 1999. **181**(8): p. 2411-21.

44. Nasser, W., J. Robert-Baudouy, and S. Reverchon, *Antagonistic effect of CRP and KdgR in the transcription control of the Erwinia chrysanthemi pectinolysis genes.* Mol. Microbiol., 1997. **26**(5): p. 1071-82.

45. Yuan, M. and C. Kendziorski, *A unified approach for simultaneous gene clustering and differential expression identification.* Biometrics, 2006. **62**(4): p. 1089-98.

46. Chen, D., et al., *A microarray analysis of the emergence of embryonic definitive hematopoiesis.* Exp Hematol, 2007. **35**(9): p. 1344-57.

47. Van Gijsegem, F., et al., *Analysis of the LacI family regulators of Erwinia chrysanthemi 3937, involvement in the bacterial phytopathogenicity.* Mol Plant Microbe Interact, 2008. **21**(11): p. 1471-81.

48. Chatterjee, A., Y. Cui, and A.K. Chatterjee, *RsmC of Erwinia carotovora subsp. carotovora negatively controls motility, extracellular protein production, and virulence by binding FlhD and modulating transcriptional activity of the master regulator, FlhDC.* J. Bacteriol., 2009. **191**(14): p. 4582-93.

49. Harris, S.J., et al., *The hexA gene of Erwinia carotovora encodes a LysR homologue and regulates motility and the expression of multiple virulence determinants.* Mol. Microbiol., 1998. **28**(4): p. 705-17.

50. Quinones, B., G. Dulla, and S.E. Lindow, *Quorum sensing regulates exopolysaccharide production, motility, and virulence in Pseudomonas syringae.* Mol Plant Microbe Interact, 2005. **18**(7): p. 682-93.

51. Shih, Y.L., et al., *The hexY genes of Erwinia carotovora ssp. carotovora and ssp. atroseptica encode novel proteins that regulate virulence and motility co-ordinately.* Environ. Microbiol., 1999. **1**(6): p. 535-47.

52. Soscia, C., et al., *Cross talk between type III secretion and flagellar assembly systems in Pseudomonas aeruginosa.* J Bacteriol, 2007. **189**(8): p. 3124-32.

CHAPTER 4: Gluconate metabolism is required for virulence of the soft-rot pathogen *Pectobacterium carotovorum*

ABSTRACT:

P. carotovorum is a ubiquitous soft rot pathogen that uses global virulence regulators to coordinate pathogenesis in response to undefined environmental conditions. This chapter discusses metabolic and nutrient signals that influence the global network of virulence regulators in *P. carotovorum*, introduced in the introductory chapter. We characterize an operon in *P. carotovorum* required for gluconate metabolism and virulence. The operon contains four genes that are highly conserved among proteobacteria (initially annotated *ygbJKLM)*, one of which was previously mis-annotated as encoding a type III secreted effector, (*ygbK,* also annotated as *hopAN1).* A mutant with a deletion-insertion within this operon is unable to metabolize gluconate, a precursor for the pentose phosphate pathway. The mutant exhibits attenuated growth on the leaves of its host of isolation, potato, and those of *Arabidopsis thaliana*, but hyper-macerates potato tubers and is deficient in motility. Lastly, the mutant is deficient in normal regulation of KdgR and FlhD, global virulence regulators that are responsive to cell wall pectin breakdown products and largely undefined environmental signals, respectively. The mis-regulation of virulence via known global virulence regulators in our *ygbJ-M* operon mutant suggests a role for host-derived catabolic intermediates in *P. carotovorum* pathogenesis. We rename this operon in *P. carotovorum vgu*ABCD for *v*irulence and *glu*conate metabolism.

INTRODUCTION

Pectobacterium carotovorum subsp. *carotovorum* is a plant-associated *Enterobacteriaceae* family member found worldwide in surface waters, soil, carrier invertebrates, as well as plant hosts [1-4]. *P. cartovorum* can infect a range of plants to cause soft-rot disease and is responsible for significant economic losses in potato production each year. The strain *Pcc* WPP14, which we discuss here, was isolated from the irrigation pond of a Wisconsin state potato farm suffering from a rot outbreak after a hailstorm. *P. carotovorum* can infect a host plant by multiple routes and can elicit disease on leaves and stems, as well as in tubers. Soft-rot outbreaks are generally triggered by environmental factors such as rain and/or hot weather and can strike during tuber storage leading to total crop loss.

Pectobacterium are often described as brute-force pathogens as their virulence strategy relies heavily on Plant Cell Wall Degrading Enzymes (PCWDE), which are secreted via a Type II secretion system (T2SS) [5-7]. Cellulases, pectate-lyases, and polygalacturonase are responsible for the characteristic rotting symptoms of infection by *P. carotovorum* and, consequently, its necrotrophic life-history [8]. However, *P. carotovorum* uses an array of virulence determinants in addition to the T2SS [9], including antibiotics, metalloproteases, adhesins, and a Type III Secretion System (T3SS), which are all tightly controlled by a well-

studied network of regulators (reviewed in [9-16]). For example, at least three regulators manipulate the production of PCWDE, KdgR, HexA, and RsmA.

One key regulator is KdgR, a repressor that blocks transcription of PCWDE genes in the absence of pectin-breakdown products [17, 18]. KdgR is a rare example where we understand how known extracellular signals directly influence the characterized signaling cascades [19]. The actions of KdgR feed into the Rsm post-transcriptional regulation system [20], a central component of virulence regulation modules in *P. carotovurom*, which is a system homologous to the Csr system in *E. coli* [21, 22]. KdgR negatively regulates *rsmB*, a functional RNA that sequesters RsmA. RsmA binds to and regulates translation of target transcripts [23], including PCWDE gene transcripts. In the presence of pectin breakdown products, KdgR de-represses PCWDE genes and *rsmB* expression, leading to PCWDE transcription and sequestration of RsmA by *rsmB* and consequent PCWDE translation. However, the Rsm system is influenced by multiple regulators including the acyl homoserine lactone receptor ExpR [24, 25], and regulates multiple virulence factors [26] including the T3SS and toxins in addition to the PCWDE [27, 28]. Thus a complex network of concerted and antagonistic regulators determines virulence expression in the variety of environments and host tissues in which *P. carotovorum* is found.

Presumably, several environmental conditions are monitored in order to transition from an environmental microbe in soil or surface water to the leaf, stem, or tuber

of a host plant and progress through stages of infection [29-32]. Just as KdgR is responsive to pectin-breakdown products, many of the other virulence regulators that feed into the Rsm system are responsive to environmental cues that have yet to be identified. The FlhDC hexomeric complex is another key virulence regulator that acts downstream of environmental signals and upstream of the Rsm system [33]. Considered the master regulator of motility, FlhDC interacts with multiple virulence regulators, including the two-component system, GacA/S, and a LysR-like transcriptional regulator, HexA [34, 35]. FlhDC and HexA are both thought to be responsive to unknown extracellular environmental factors, and HexA, in turn, regulates *rsmB* expression. The sensor kinase GacS is thought to detect cell density, growth phase and other undetermined environmental factors, while its cognate response regulator GacA is itself regulated by medium composition and growth phase [33, 36, 37]. GacA regulates toxins, PCWDE and *rsmB* downstream of signaling from GacS [38]. In addition to regulating HexA and GacA/S, FlhDC in *E. coli* is responsible for regulating a shift to aerobic respiration and genes involved in the Entner-Doudoroff pathway, an alternative glycolysis pathway [39, 40]. Thus, the activity of FlhDC not only provides a clear regulatory link between virulence and motility [41], but metabolism and environmental conditions as well.

In this chapter, we characterize a *P. carotovorum* operon required for gluconate metabolism that modulates the expression of key virulence regulators and affects virulence in plants. In 2002, Boch *et al.* employed an IVET screen in *P. syringae*

for T3SS virulence determinants, in which they identified a gene they designated *ipx53* [42]. *ipx53* was considered a potential type III effector for a short time, and given the name *HopAN1*. Unlike other T3SS effectors, this gene is highly conserved in enterobacteria and plant pathogens [43-45]. The *hopAN1* homolog of *Pcc* WPP14 exists in an operon of four genes encoding, in order: an oxidoreducatase, *hopAN1*, a class II aldolase and an isomerase [46] (Fig. S1). The operon is highly conserved in enterobacteria and the four genes were named *ygbJKL* and *M* in *E. coli.* The designation *ygb* is the temporary name for genes of unknown function [47]. We have therefore named the *Pcc* operon *vguABC* and *D* for virulence and gluconate. The *vgu* operon is required for gluconate metabolism. In *P. carotovorum,* glucose may be oxidized to 2,5-diketogluconate and subsequently reduced to gluconate which can be metabolized through the pentose phosphate pathway as it is in related species [48]. The predicted functions of the *vgu* operon-encoded proteins suggest a pathway similar to the pentose phosphate pathway. In the *Pcc* WPP14 *vgu* mutant described here, key virulence regulators *flhD*, *kdgR*, and *rsmB* are mis-expressed during infections. Thus, we provide evidence suggesting that gluconate catabolic products affect virulence determinants by modulating known regulators.

MATERIALS AND METHODS

Bacterial strains, plasmids and media: Bacterial strains were maintained in Lysogeny Broth (LB) agar or 2xYT broth, containing appropriate antibiotics,

Sambrook and Russell, 2001. The Minimal Media for *hrpL* repression was described in [49]. Tuber extract media was made by autoclaving 200 g of chopped Yukon gold potatoes with 200 mL of water. After autoclaving the water-potato slurry was centrifuged at 12,000Xg for 30 min. The supernatant was diluted 1:5 in M9 Minimal media lacking a carbon source. When required, antibiotics and drugs were supplemented at the following concentrations: ampicillin [39] 100 μg/mL; chloramphenicol (Cm), 30 μg/mL; gentamycin (Gm), 25 μg/mL; kanamycin (Km), 30 μg/mL; rifampicin (Rif), 100 μg/mL; spectinomycin (Spec), 50 μg/mL; and tetracycline (Tet), 5 μg/mL. Media was solidified with 1.5% wt/vol BD Bacto-agar. Cultures were grown at 28 ºC and, if in liquid, shaken at 250 rpm)

DNA manipulations

Standard procedures were used for: plasmid and chromosomal DNA isolation; electroporation; restriction endonuclease digestions; ligations; gel electrophoresis; and tri-parental mating. Enzymes were obtained from either New England Biolabs or Invitrogen. Nucleotide sequences were determined by the UNC-CH Genome Analysis Facility and sequences were analyzed using BLAST.

Construction of the *P. carotovorum vgu* operon mutant

To replace the *vgu* operon, we used splicing over-lap extension (SOE) PCR [50] to create a construct with a Cm cassette flanked by 1 kb of sequence surrounding *vguB* on each side. Briefly, a 1 kb region upstream of *vguB* was

amplified using SOE primers containing a *Sal*I site for the insertion of a Cm cassette on the 3' end, and a 1 kb downstream region was amplified similarly with the *Sal*I site on the 5' end with SOE primers (shown as primers 1 and 2 in Table S1). These amplicons were then fused together in a two-step SOE PCR reaction and the SOE product was cloned into pCR2.1 Topo TA cloning vector (amp^R) (Invitrogen) to make pCR2.1SOE*vgu*, which was then digested with *Sal*I. The Cm cassette was amplified from pKD3 using modified primers from [51] to contain *Sal*I sites on both ends (primers 3 and 4 in Table S1) The Cm cassette was then digested with *Sal*I and ligated to the linear pCR2.1SOE*vgu* to make pCR2.1Δ*vgu.* This plasmid was introduced into *P. carotovorum* WPP14 by electroporation. The resulting *Pcc* WPP14::pCR2.1Δ*vgu* strain was then grown in potassium phosphate buffer supplemented with Cm, but not Amp. After over-night growth, the cultures were transferred to 2XYT supplemented with Cm and passaged for three days, at which point the cultures were replica plated on LB plates that contained either Cm and Amp, or just Cm, to identify strains that had lost the plasmid but undergone double-homologous recombination for marker exchange. The resulting deletion-insertion mutants were verified by PCR analysis and sequencing across the entire four-gene operon.

Table S1. List of primers used.

#	Designation	Sequence, 5'-3'
1	opkoupfwd	ACAAACAGACATAAACACCCCCCC
2	opkuprev	GGGAACAGCAATATTTTTCCGCGATAACGTCGACCGTGGTTGCCTTGTTAAAATTAGTG
3	opkodownfw	CACTAATTTTAACAAGGCAACCACGGTCGACGTTATCGCGGAAAAATATTGCTGTTCCC
4	opkodownrev	CCGCCTCCCCATTGTGAAATTTTACGCT
5	1180pro Fwd	CACCGCTCAGCAGCAGGCCGTCCTGCTCCAACTTTTGCACATCGCGCCG
6	3244 Rev	TTATTTCTTTAGATAGGGTTTAGCCCATCCC

Supplemental Table 1. List of primers used.

Complementation of *vgu* operon mutant

The four-gene operon and a 500 bp region encompassing the operon promoter region were amplified with *pfx* polymerase (Invitrogen) with primers 5 and 6 (Table S1): Gel electrophoresis with a 1% (wt/vol) agarose (Invitrogen ultrapure agarose) was used to view and isolate the resulting amplicon. Gel extraction buffers and protocols were supplied from Qiagen Gel extraction kit (Qiagen). The isolated amplicon was then incubated in the presence of Taq polymerase (Invitrogen) and PCR reaction mix without primers for five minutes at 70 °C in order to add overhanging As on the blunt amplicon. The amplicon was then cloned into pCR2.1 TA cloning vector (Invitrogen TA cloning kit) according to manufacturer's protocol. The resulting plasmid was then conjugated into the *P. carotovorum vgu* operon mutant via tri-parental mating with *E. coli* helper strain pRK2013. The resulting complementation strain was isolated on Cm and Kan plates and verified by PCR from vector borne M13 primers (Invitrogen).

***A. thaliana* and Potato Leaf Infection Assay**

Cultures were grown overnight in 2XYT media, washed twice in 10 mM $MgCl_2$, and resuspended at 10^4 CFU/mL in 10 mM $MgCl_2$. Cultures were hand infiltrated with a needle-less 1 ml syrinage into the leaves of either 4-5 week old Yukon gold potatoes or 4-5 week old *A. thaliana.* At subsequent time points, infected leaves were cored with a gauge 3 (6 mm diameter) cork borer and the tissue ground in the presence of 10 mM $MgCl_2$. The CFU/mL of present bacteria was quantified by titration and plating on LB plates with the appropriate antibiotics.

Tuber Maceration Assay

Tubers were injected with 10 μL of resuspended at a concentration of 10^8 CFU/mL. Bacteria were inoculated into 15 mm holes in the tuber, made with a pipette tip. The infected tubers were then placed in a plastic bag, which was sealed and kept at 28 ºC for 5 days. After 5 days, the soft, macerated tissue surrounding each injection sight was carefully scooped out using a metal spatula and weighed. Each experiment contained 10 internal replicates.

Motility Assays

Motility media recipes were described in [52]. Swimming motility media contained Tryptone (10 g/L), NaCl (5 g/L) and Bacto-Agar (BD) at 0.3% wt/vol. Swarming motility media contained Nutrient Broth (8 g/L), and Bacto-Agar at 0.5% wt/vol. Twitching motility media contained Tryptone (10 g/L), Yeast Extract (5 g/L), NaCl (10 g/L) and Bacto-agar at 1% wt/vol. Each motility plate contained 50 mLs of motility media with appropriate antibiotics. Plates were inoculated with 10^6 CFU of bacteria and incubated for 18 hours at room temperature, at which point the diameter of the colony was measured.

RNA Extraction

From leaves: *A.* thaliana or potato leaves were infiltrated with wild-type *Pcc* WPP14 cultures at an OD_{600} of 1.0 (approximately $4x10^8$ CFU/mL) using a needle-less syringe. Two and a half hours post infection, 10 leaves were

collected, ground in the presence of Qiagen Protect Bacteria and homogenized over a Qiashredder column before RNA was extracted using the RNeasy kit (Qiagen).

From tuber extract media: Cultures were grown in rich media overnight, centrifuged and washed in 10 mM $MgCl_2$. Bacteria was resuspended at an OD_{600} of 0.5 ($2x10^8$ CFU/mL) and grown for 3 hrs, shaking at 28 ºC. Bacteria were then harvested in the presence of Protect Bacteria and RNA was extracted with the RNeasy mini kit (Qiagen) However, the protocol was modified to include an RNA precipitation step with LiCl after lysis but before the ethanol precipitation and RNA being bound to the column. This step allowed for the removal of excessive sugars present in the tuber extract media.

qRT-PCR

RNA was reverse transcribed into cDNA using the Ambion RETROscript kit. Briefly, 2 ug of RNA was primed with random decamers and reverse transcribed with the MMLV-RT. cDNA was then diluted and relative quantities of specific transcripts were determined using SYBR Green RT-PCR Reagents. Fluorescence of double-stranded DNA was measured by a DNA Engine, Opticon 2 Continuous Fluorescence Detector (MJ Research) and values were analyzed using Opticon Monitor 3 software (MJ Research). Relative gene expression was determined by normalizing to a *P. carotovorum* housekeeping gene, *ffh,* which

encodes a signal recognition particle protein, and was identified as an optimally stable control gene for qRT-PCR analysis by [53].

Carbon Source Utilization Assays

Biolog GN2 plates, which contain 95 discrete carbon sources, were inoculated with *P. carotovorum Pcc* WPP14, the *vgu* operon mutant, and the complemented *vgu operon* mutant. Briefly, cells were grown in LB media at 28°C for 24 hours, washed twice in 10 mM $MgCl_2$, and resuspended in 10 mM $MgCl_2$ to a final concentration of 0.4 OD_{600}. Aliquots of 150 µL were added to each well. The plates were sealed and allowed to incubate statically at room temperature for 24 hours. Each plate contained a negative control that lacked a carbon source to ensure that metabolic activity of stored carbon reserves was not being measured. In response to respiration, tetrazolium in each well can become reduced to produce a distinctive purple color indicating the oxidization of the available carbon source Carbon utilization was additionally tested by growth in M9 Salts and 100 mM of a specific carbon source. Bacteria were grown in liquid 2XYT culture over night, washed in M9 salts two times and resuspended at OD_{600} of 0.04 in the M9 Salts with a carbon source. The culture was then incubated for 18-20 hours at 28°C, while shaking, and the OD_{600} was measured again.

Gas Chromatography coupled to Mass Spectrometry

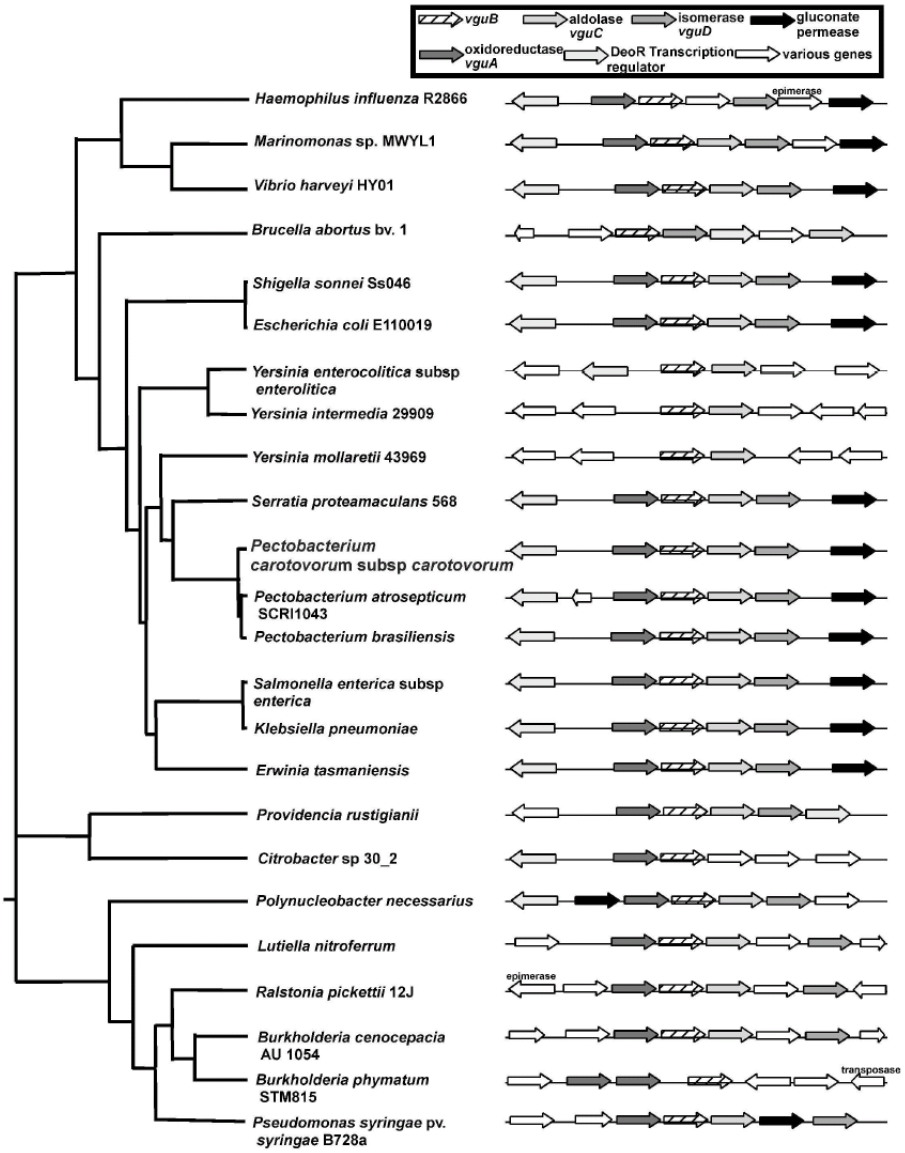
Preparation and methods for GC-MS analysis was as described [54]. Cultures of wild type and the *vgu* operon mutant were grown from single colonies in LB media overnight. The next day, over-night cultures were used to inoculate 100 mLs of potato extract media and grown to an OD_{600} of 0.4 in approximately 4 hours, shaking at 250 r.p.m at 28 °C. Cellular metabolism was immediately quenched by addition of methanol at -45 °C. Intracellular metabolites were extracted with chloroform as described in [55]. At -45 °C chloroform was added to the methanol/water mixture to break the cell walls and denature enzymes. The water/methanol phase was then lyophilized and derivatized with 10 uL of 56 mg/mL ethoxyamine hydrochloride solution in pyridine and 20 uL of pyridine. The samples were then incubated at 40 °C for 90 min. The samples were then silylated for 50 min at 40 °C with 70 uL of N-Methyl-N-trimethylsilyltrifluoroacetamide (MSTFA). The derivatized extracts were prepared for GC-MS analysis by the addition of ethyl acetate as a carrier and filtering in the presence of acetyl nitrile to ensure the absence of metal ions. Samples were analyzed with an Agilent 6850 gas chromatograph coupled to an Agilent 5973 mass selective detector. The 5 uL aliquots of extract were injected into a capillary column (30 m_.25 mm i.d., 0.25 m film thickness) at 250 °C. The initial temperature of the gas chromatograph was 50 °C and held for three min before ramping to 250 °C at 10 °C per minute. Helium was used as a carrier gas. Detection was achieved using MS detection in electron impact mode and full scan monitoring mode (m/z 10_550). Data was analyzed using ChemStation (Enhanced Software).

RESULTS

The *vguABCD* operon is widely distributed in proteobacteria. The structure of the *vguA-D* operon in *P. carotovorum* is generally conserved in enterobacteria. A similar operon also exists in more distantly related plant pathogenic bacteria although one or two of the ORFs are missing in some species. The conservation of the operon structure in select proteobacteria is illustrated in Supplemental Figure S1. Many of the chromosome regions are flanked by other genes predicted to be involved in sugar metabolism, such as epimerases and decarboxylases. The conserved genomic structure of this operon suggests its products could be involved in a metabolic function.

The *vgu* operon mutant has significantly attenuated growth on potato and Arabidopsis leaves. To determine if the *vgu* operon has a role in virulence, we constructed a deletion-insertion mutation that replaces *vguB* with a chloramphenicol cassette, which has a polar effect and disrupts the expression of both the up and downstream overlapping genes in the operon as determined by RT-PCR. This mutation does not affect expression of the downstream gluconate permease gene, *gntP* (data not shown). The mutation was confirmed by PCR and re-sequencing of the entire *vgu* operon. The *vgu* operon mutant had the

Figure S1.



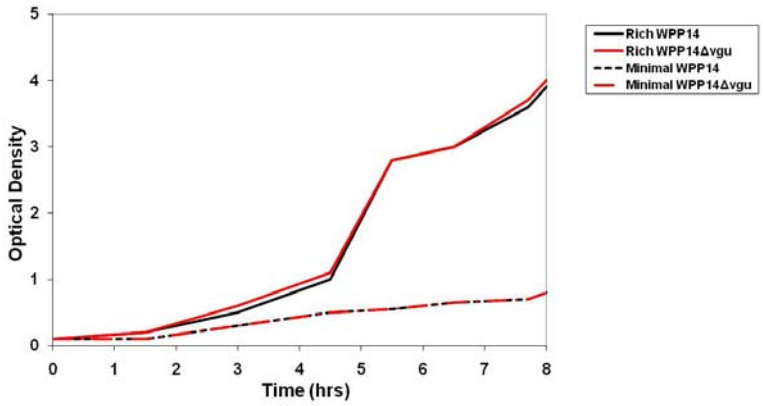
Supplemental Figure 1. *vgu* operon organization is conserved in enterobacteria. Phylogenetic tree constructed using maximum-likelihood estimation with a molecular clock of *vguB* homologs from select species, which represent the range of *vgu* distribution. Coupled to the tree is a cartoon depiction

of the chromosome region in which the *vgu* homolog lies. The structure of the *vguABCD* operon in *P. carotovorum* is generally conserved in species with the closest *vguB* homologs. The operon is often flanked by a DeoR-like transcription regulator on one side and a gluconate permease on the other. However, some species lack the complete operon, such as *Yersinia enterocolitica, Yersinia intermedia,* and *Yersinia mollaretti* , which appear to only carry the *vguB* homolog and the *vguC* putative aldolase in the region. Interestingly, *Burkholderia phymatum* contains two copies of the *vguA* oxidoreductase gene, but lacks the *vguC* and *vguD* genes, two genes at the 3' side of the operon. Nearly all of the chromosome regions are flanked by other genes predicted to be involved in sugar metabolism, such as epimerases and decarboxylases.

---

same growth rate in both rich and minimal media types as wild type (Figure S2). To determine if the *vgu* operon is involved in virulence, we compared growth of the wild type to that of the *vgu* operon mutant in its native host, potato (Fig. 1A), as well as a model host, *A. thaliana* (Fig. 1B). In both potato and *A. thaliana* leaves, *Pcc* WPP14 grew to 10^8 CFU/mL by three days post infection. In potato leaves, the *vgu* operon mutant grew three logs less than *Pcc* WPP14 by three days post infection, illustrating a severe attenuation in virulence. Similarly, in *A. thaliana* leaves, the *vgu* operon mutant grew two logs less than *Pcc* WPP14 by three days post infiltration on leaves of both hosts. The *vgu* operon mutant failed

Figure S2.



Supplemental Figure 2. Growth of the *vgu* operon mutant in rich (2XYT) and minimal media. The optical density$_{600}$ of both wild type (strain WPP14) and the *vgu* operon mutant were both started in liquid cultures at 0.1 OD_{600} and monitored over the course of 8 hours for growth.

---

to produce maceration symptoms typically seen during *P. carotovorum* infection on leaves of both hosts. These begin as a rotting lesion spreading out from the site of infiltration over the course of the growth experiment. Both attenuation of growth and lack of maceration mutant phenotypes were rescued with the *vgu* operon cloned with its native promoter in the entry clone pCR2.1 and mated into the *vgu* operon mutant strain (Fig. 1).

Figure 1

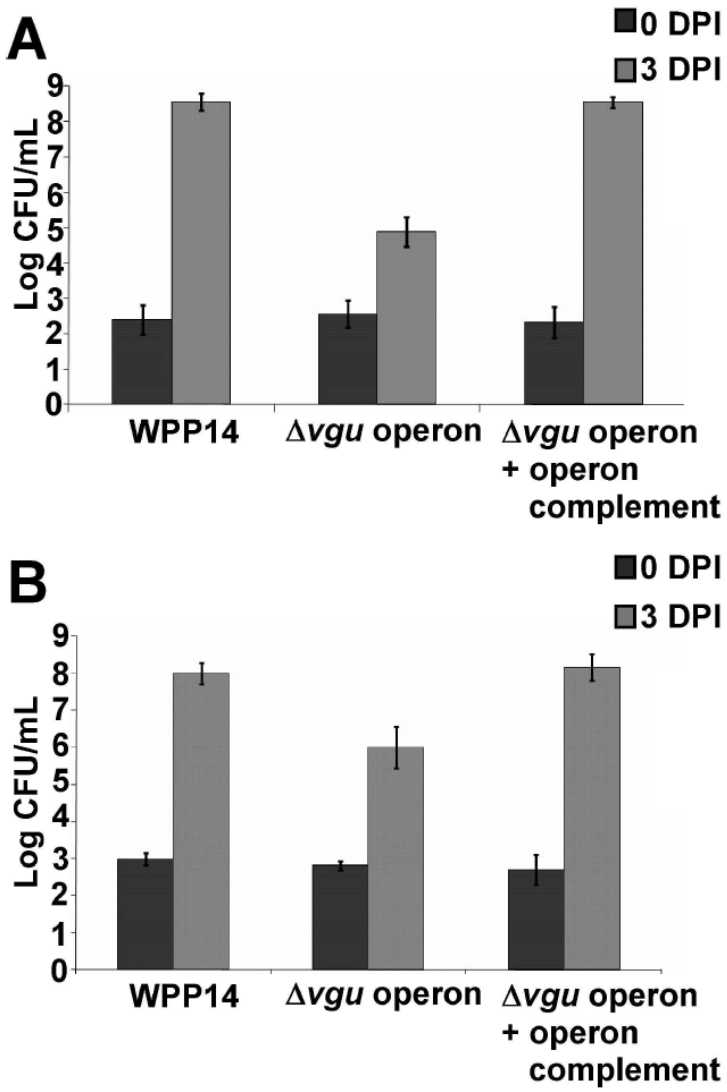


Figure 1. ***VguA-D* mutant is defective in growth on leaves**. Potato (A) and *A. thaliana* (B) hand infiltration of *Pcc* WPP14, the *vgu* operon, or the *vgu* operon mutant carrying pCR2.1:*vgu* operon with 10^4 CFU/mL initial inoculum, into leaves of 4-5 week old plants. Experiment has four internal replicates and was repeated three times; error bars indicate two times standard error.

Figure 2

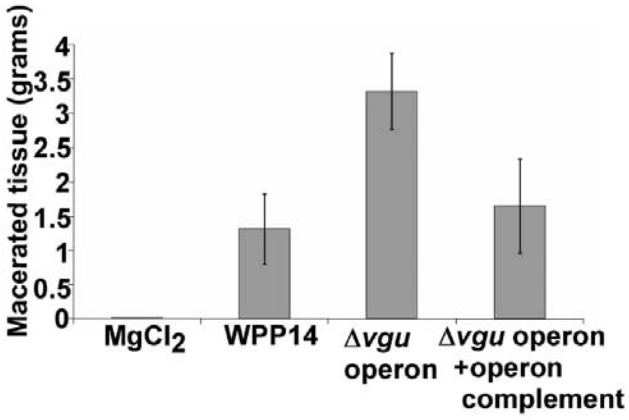


Figure 2. **The *vgu* operon mutant hypermacerates infected potato tubers.** Potato tubers were inoculated with either 10 mM $MgCl_2$, *Pcc* WPP14, the *vgu* operon mutant, or this mutant mutant carrying pCR2.1:*vgu* operon. 10^6 CFU were injected into a 1.5 mm hole and tubers were incubated at 28 ^oC for five days. Experiment has ten internal replicates, error bars are two times standard error and was repeated twice.

The *vgu* operon mutant hyper-macerates potato tubers compared to wild type. To determine if the *vgu* operon mutant was also defective in tuber maceration, we compared wild type maceration to that of the *vgu* operon mutant. Each tuber was infected with three strains: *Pcc* WPP14, the *vgu* operon mutant, and the complemented *vgu* operon mutant, as well as a negative control for maceration, 10 mM $MgCl_2$ buffer. The wild type and the *vgu* operon mutant carrying the complementation clone both macerated just over 1 gram of potato tuber in 5 days, while the *vgu* operon mutant macerated over 3 grams of potato

tuber (Fig. 2). The concentration of bacteria per mg of tissue was approximately $1X10^9$ for both the mutant and wild type. This indicated that the *vgu* operon mutant is hyper-macerating, suggesting a mis-regulation of the expression of plant cell wall degrading enzymes that characterize soft rotting diseases. This hyper-virulence in the tuber was unexpected, given the attenuated virulence phenotypes on leaves reported above.

The *vgu* mutant is deficient in swimming motility. Previous studies of *P. carotovorum* and other phytopathogens have demonstrated co-regulation of motility modes during pathogenesis. Therefore, we compared the *vgu* operon mutant to the wild type on twitching, swarming and swimming motility agar plates.

Figure 3

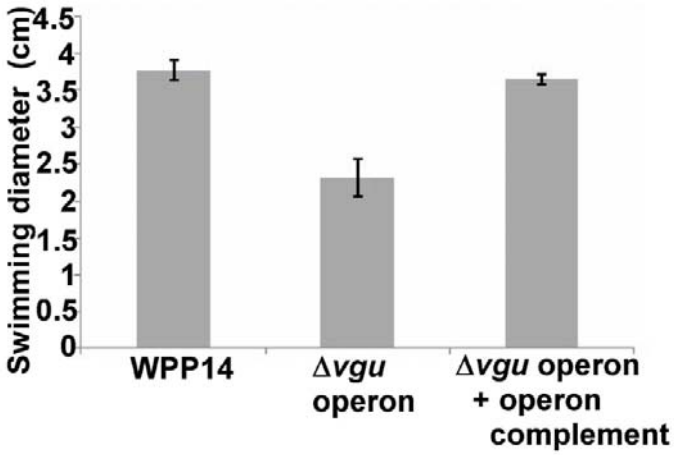


Figure 3. The *vgu* operon mutant is defective in swimming motility. Swimming motility plates were inoculated with 10^6 CFU into the center of the plate, each with 50 mL of motility agar and appropriate antibiotics. Plates were

grown for 18 hrs at room temperature. Experiment has four internal replicates and was repeated three times, error bars indicate two times standard error.

The *vgu* operon mutant was able to swarm and twitch like wild type *Pcc* WPP14, however, the *vgu* operon mutant exhibited significant reduction in swimming plates. On motility plates, the wild type and the *vgu* mutant carrying the complementation clone swam almost twice the diameter of the *vgu* operon mutant (Fig. 3). This indicates that the *vgu* operon mutant is also unable to properly swim and / or perform chemotaxis.

The *vgu* operon is neither HrpL-regulated, nor up-regulated in leaves. *vguB* has homology to the *P. syringae* gene annotated as *hopAN1*. The name HopAN1 indicates that the gene potentially encodes a type III effector protein [56]. We determined if HrpL, the alternative sigma factor responsible for T3SS regulation in *P.* carotovorum [57], regulates *vguB* transcription in *Pcc* WPP14. Cultures of *Pcc* WPP14 containing the native *hrpL* ORF cloned into pCF430, where *hrpL* is under the control of the arabinose-inducible promoter [49], were grown with or without 200 mM arabinose in *hrpL*-repressing media. qRT-PCR was used to assess the relative expression of *vgu* and a known HrpL-dependent T3SS effector, *dspE* (Fig. 4A). In the presence of over-expressed *hrpL*, *dspE* expression increased over 200-fold compared to the absence of *hrpL* over-

expression. *vgu* expression, however, was not significantly altered by *hrpL* over-expression. By contrast to the *P. syringae hopAN1* gene, *vgu* expression is not

Figure 4

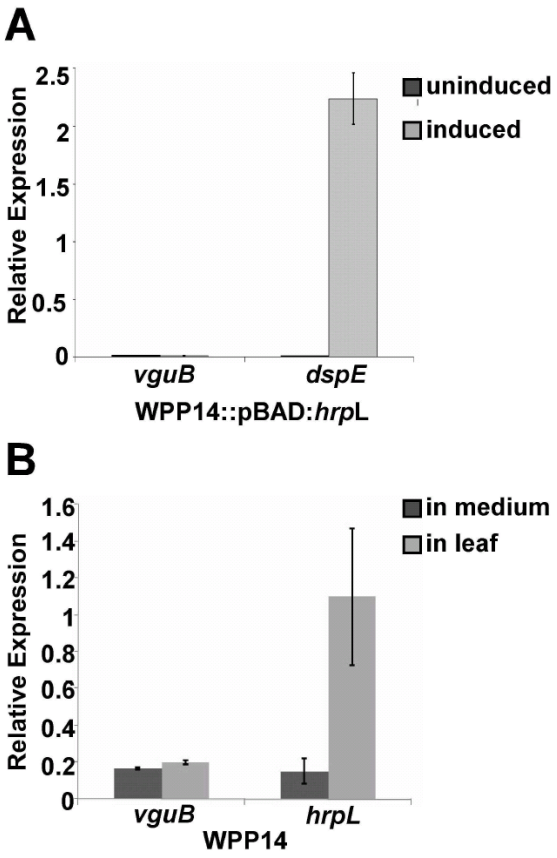


Figure 4. Expression of the *vgu* operon is independent of T3SS regulation. (A) qRT-PCR data showing relative expression of *vgu* and *dspE* expressed in *Pcc* WPP14 carrying pCF430 with $hrpL_{WPP14}$. Cultures were grown in *hrp*-repressing minimal media and *hrpL* expression was induced with 200 mM Arabinose. (B) qRT-PCR data showing relative expression of *vgu* and *hrpL* expressed in *Pcc* WPP14. Strains were grown either in *hrp*-repressing minimal

media overnight or infiltrated into *A. thaliana* leaves at 10^8 CFU/mL with RNA extraction 2.5 hpi**.** Expression is relative to the *ffh* housekeeping gene.

---

induced in infected plants. *W*e infiltrated *A. thaliana* leaves with wild-type *Pcc* WPP14 cultures with approximately $4x10^8$ CFU/mL. RNA was extracted 2.5 hours post-infection. Simultaneously, a *Pcc* WPP14 culture was grown in *hrp*-repressing minimal media to the same concentration as the infiltrated culture and RNA was extracted. qRT-PCR performed using RNA from infected leaves or culture indicated that *hrpL* transcription increased ~7-fold in leaves compared to expression in the *hrp*-repressing minimal media (Fig. 4B). The expression of *vguB*, however, was not altered in leaves compared to minimal media, and showed low-level expression in both conditions.

The *vgu* operon mutant exhibits altered expression of key virulence regulators in leaf infections and in the presence of tuber extract. Our *in vivo* virulence assays indicated that virulence of the *vgu* operon mutant was attenuated in the leaf, yet over-active in the tuber. To address the virulence regulation changes that could account for this observation, we used qRT-PCR to assess the expression level of known virulence regulators in both conditions. In leaves, the mutant exhibited higher mRNA levels of the pectin-breakdown responsive repressor, *kdgR,* and the T3SS helper protein *hrpN* than wild type (Fig. 5A). In the presence of tuber extract, we found that the master regulator of motility, *flhD* [33] and the LysR-like transcription regulator, *hexA* [34] were

severely down-regulated in the *vgu* mutant compared to wild type (Fig. 5B). Though weakly expressed relative to the *ffh* control, the key virulence regulator *rsmB* was also down regulated 8-fold in the *vgu* operon mutant compared to the complementation strain in the presence of tuber extract (Fig. 5B). Following leaf infection, *flhD* and *rsmB* were both expressed at very low levels in the *vgu* operon mutant (Fig. 5A).

Figure 5

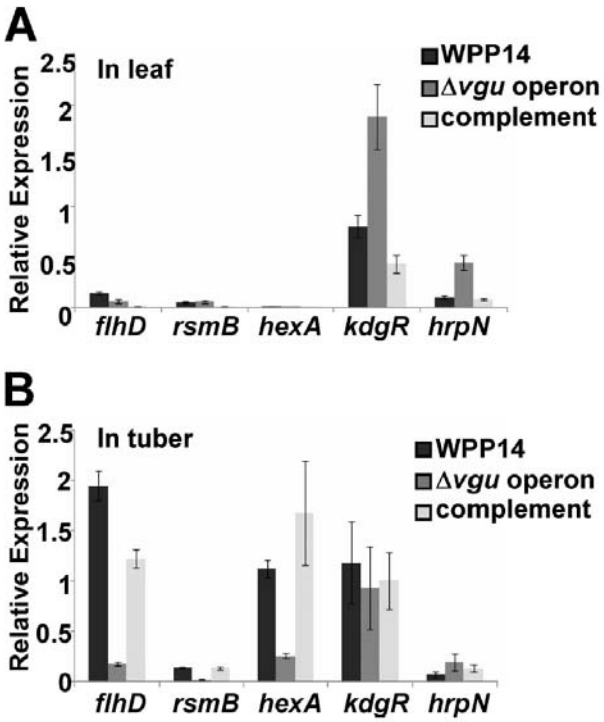


Figure 5. The *vgu* operon is required for normal expression of key virulence regulators in leaf and tuber extract. qRT-PCR data showing relative expression of key virulence regulators and determinants in, *Pcc* WPP14, the *vgu* operon, or the *vgu* operon mutant carrying pCR2.1:*vgu* operon, from leaf infection conditions after 2.5 hours and in the presence of tuber extract after three hrs of growth. Expression is relative to the *ffh* housekeeping gene.

The vgu operon is required for gluconate metabolism. To address whether the vgu operon is involved in sugar metabolism, as BLAST analysis suggests, we used a phenoarray plate to assess the metabolic phenotype of our *vgu* mutant in relation to wild type. The Biolog GN2 phenoarray plate indicated that wild type *Pcc* WPP14 was capable of using all carbon sources tested with the exception of D-gluconate (carbon sources tested are listed in Table S3)**.** The ability to grow on gluconate was rescued by the expression of the full operon *in trans* in the complementation strain. In *P. carotovorum*, gluconate is a precursor sugar for the pentose phosphate pathway [48] and the first enzyme in the *vgu* operon has similiarity to the first enzyme in the pentose phosphate pathway, 6-phosphogluconate dehydrogenase. However, the mutant was able to grow identically to wild type on carbon sources that are sugar intermediates of the pentose phosphate pathway, xylulose and ribulose. This may indicate that the defect in the *vgu* mutant is limited to early steps in the pentose phosphate pathway or that these pentose sugars can be metabolized through multiple pathways. In preliminary assays using gas chromatograph coupled to mass spectrometry to compare the metabolite profiles of wild type and the operon muntant, the mutant was deficient in the accumulation of pentose sugars including erythropentose, a pentose phosphate intermediate. It was also deficient in some hexose sugars, galactopyranuronic acid and the glucose structure glucopyanose When grown in either minimal or rich media, the *vgu* operon mutant does not have growth defects relative to wild type.

DISCUSSION

We demonstrated that the *vgu* operon is involved in gluconate metabolism and is required for proper regulation of virulence via key virulence regulators KdgR, FlhD, HexA, and the *rsm* system. Subsequently, a mutant lacking the *vgu* operon has attenuated virulence in leaves, hyper-macerated potato tubers, and is deficient in motility. The inability of the mutant to grow on gluconate and its lack of intermediate pentose sugars from gluconate metabolism suggest that the *vgu* operon encodes enzymes for gluconate metabolism. Our findings strongly indicate that gluconate metabolism is critical for virulence expression through previously identified regulators.

Gluconate metabolism in *P. carotovorum*

In environments with high glucose concentrations, some bacteriaoxidize glucose extracellularly to gluconate or 2-ketogluconate, through membrane-bound dehydrogenases. When glucose is depleted in the environment, the oxidized glucose can be transported into the cell and phosphorylated in an ATP-dependent manner, forming 6-phosphogluconate or 2-keto-6-phosphogluconate, respectively. These phosphorylated, intracellular sugars can then be used for the Entner-Dourdoroff (Liu *et al.*) pathway, an alternative to the Embden-Meyerhof-Parnas (EMP) glycolysis pathway, or the pentose phosphate shunt [48, 58, 59].

Table S2.

Carbon Sources	WPP14	Δ*vgu* operon	Complement
Acetic Acid	+	+	+
α–D-Glucose	+	+	+
α-D-Glucose-1-P	+	+	+
α-D-Lactose	+	+	+
α-Keto Butyric Acid	+	+	+
α-Keto Glutaric Acid	+	+	+
Bromosuccinic Acid	+	+	+
Cis-Acontic Acid	+	+	+
Citric Acid	+	+	+
D-cellulobiose	+	+	+
Dextrin	+	+	+
D-Fructose	+	+	+
D-Galactose	+	+	+
D-Galacturonic Acid	+	+	+
→D-Gluconate	+	--	+
→ Ribulose	+	+	+
→ Xylulose	+	+	+
D-Glucose-6-P	+	+	+
D-Glucuronic Acid	+	+	+
D-L, alpha Glycerol phosphate	+	+	+
D-Mannitol	+	+	+
D-Mannose	+	+	+
D-Saccharic Acid	+	+	+
Formic Acid	+	+	+
Gentiobiose	+	+	+
Glucuronamide	+	+	+
Glycerol	+	+	+
Glycyl-L-Aspartic Acid	+	+	+
i-Erythritol	+	+	+
Inosine	+	+	+
Lactulose	+	+	+
L-Alaniamide	+	+	+
L-Arabinose	+	+	+
L-Asparagine	+	+	+
L-Aspartic Acid	+	+	+
L-Glutamic Acid	+	+	+
L-Serine	+	+	+
m-Inositol	+	+	+
N-Acetyl-D-Glucosamine	+	+	+
Succinamic Acid	+	+	+
Succinic Acid	+	+	+
Thymidine	+	+	+
Tween	+	+	+
Uridine	+	+	+

Supplemental Table 2. Growth of *Pcc* WPP14 and the *vgu* operon mutant in simple sugars as a sole carbon source.

The genome of *P. carotovorum* encodes genes for ketoaldonic acid metabolism and may oxidize extracellular glucose to 2,5-diketogluconate, which can be reduced to gluconate in two separate pathways [48]. However, once in the cell,

6-phosphogluconate may solely be metabolized through the pentose phosphate pathway, since enzymes for the ED pathway were not found in significant levels in a related species [48]. Generally, if 6-phosphogluconate were to be metabolized through the pentose phosphate shunt, it would first be converted to ribulose 5-phosphate by the enzyme 6-phosphogluconate dehydrogenase. The first gene in the *vgu* operon has high similarity to 6-phosphogluconate dehydrogenase. Fitting with this prediction, preliminary GC-MS data indicates that ribulose does not accumulate in the mutant.

The ribulose 5-phosphate resulting from the first step of the pathway, is further re-arranged by transaldolases and isomerases to 6-glucose-phosphate and glyceraldehyde 3-phosphate. Although *vguB,* the second gene in the operon, does not have similarity to any other protein with a known function, the last two operon members encode a putative aldolase and isomerase. It should be noted that the *P. catorovorum* genome contains genes that are predicted to encode all other enzymes typically involved in the pentose phosphate pathway, and also encodes two gluconate transporters, one of which is encoded just downstream of the *vgu* operon. However, these other enzymes are dispersed around the genome and many are neither within an operon nor clustered with other metabolic genes. Our observation that the *vgu* operon mutant is incapable of metabolizing gluconate suggests that these other predicted enzymes are either not expressed in the conditions we tested or are involved in different pathways. In preliminary assays using gas chromatography coupled to mass spectrometry

to compare the metabolite profiles of wild type and the operon muntant, the mutant was deficient in the accumulation of pentose sugars including erythropentose, a pentose phosphate intermediate. Therefore, based on our evidence that the *vgu* operon is required for gluconate metabolism and the operon member's homology to enzymes in the pathway, we predict that the *vgu* operon is involved in a pathway similar to the pentose phosphate pathway.

The pentose phosphate pathway is also a main method for generating reducing power, via the production of NADPH, which is mainly used to prevent oxidative stress. Plant hosts produce reactive oxygen species as a defense against invasion [60]. However, in oxidative stress assays using hydrogen peroxide, the operon mutant was not more sensitive than wild type to oxidative stress. Additionally, the operon mutant did not produce different pH conditions in macerated tuber tissue or in medium, suggesting that the virulence phenotypes of the mutant are not due to stress conditions in the host.

Gluconate metabolism affects the expression of key regulators

Previous studies identified a complex network of transcriptional and post-transcriptional virulence regulators (for reviews, see [10, 61, 62], which regulate the T2SS and macerating enzymes, the T3SS, motility, quorum sensing, and antibiotic production. In leaves, the *vgu* operon mutant over-expresses *kdgR*, the transcriptional repressor known to regulate *hrpN* and genes encoding macerating enzymes [17]. KdgR activity is responsive to pectin breakdown products and

KdgR releases from target promoters after binding to catabolic intermediates of pectin degradation, such as 2-keto-3-deoxygluconate [17, 63, 64]. Thus in the presence of pectin and its catabolic intermediates, such as polygalacturonate (PGA), KdgR's repressor activity is blocked and KdgR targets, such as genes encoding macerating enzymes and *hrpN*, are expressed. Previous work by Nasser *et al.* indicated that the presence of gluconate, and presumably the catabolic intermediates of gluconate, inhibits the production of macerating enzymes in the presence of inducers PGA and galacturonate. Thus, in the *vgu* mutant deficient in gluconate metabolism, we see hyper-maceration and an increase in *hrpN* expression. HrpN is a T3SS-associated helper protein, known to induce plant defense responses. Over-expression of HrpN would induce plant defense responses sufficient to explain the *vgu* mutant's dramatic loss of virulence on leaves. A model for this activity is shown in Figure 6A, in which KdgR's repressor activity is blocked by pectin breakdown products, but sustained with the addition of gluconate. From our data and the previously published data by Nasser *et al.*, we hypothesize that the catabolic pathways for pectin and gluconate are antagonistic. The increase that we see in *kdgR* expression in the *vgu* mutant is presumably due to auto-regulation of the repressor, however there is no previously known data on *kdgR* regulation.

In tuber extract media, the *vgu* operon mutant shows no relative changes in *kdgR* or *hrpN* expression. However, the *vgu* mutant does show down-regulation

of *flhD, hexA* and *rsmB* in tuber extract compared to wild type. These genes are

Figure 6.

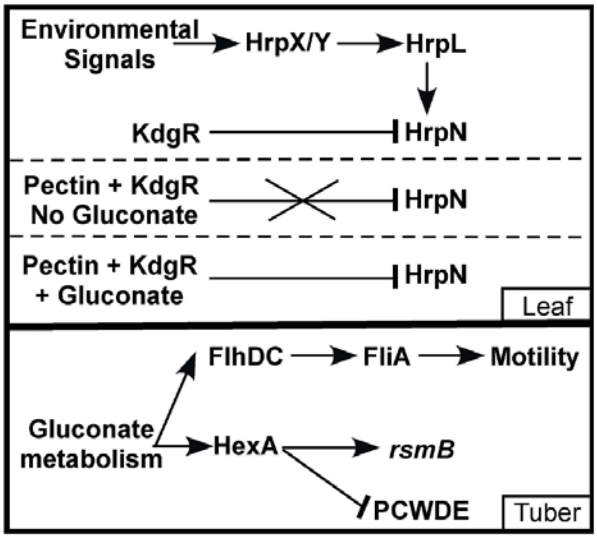


Figure 6. Hypothetical model of the *vgu* operon's role in virulence via *Flh*DC regulation in potato tubers and *kdg*R regulation in potato or *A. thaliana* leaf infection.

---

lowly expressed in both wild type and mutant during leaf infection.. The heterodimeric complex, FlhDC, was first identified as the master regulator of motility [65] in *E. coli,* and regulates the switch from aerobic respiration to anaerobic respiration [39]. FlhDC interacts with HexA, a LysR-like transcription factor, as well as the Rsm system [66]. HexA directly binds the promoters of genes encoding extracellular enzymes and its down-regulation was previously reported to cause a hyper-macerating phenotype [34]. Likewise, down-regulation of *flhD*, the master regulator of the flagellar system, in the *vgu* operon mutant can

explain the deficiency in motility that we observed in the *vgu* operon mutant on motility plates. Previous studies, however, have indicated that FlhD and HexA activity are antagonistic [66]. This begs an explanation for how both could be down-regulated simultaneously in our *vgu* operon mutant. The answer may lie in unresolved environmental factors that can putatively influence the expression of both *flhD* and *hexA*. For example, *flhD* is regulated by FliT, which is responsive to unknown environmental factors [66]. Similarly, *hexA* expression is also mediated by unknown environmental factors through an unknown mechanism [66]. In Figure 6B, we propose a model for the influence of the *vgu* operon and inferred gluconate catabolic intermediates on these key regulators during tuber infection that correlates the observed phenotypes, expression data and published regulatory network information [33, 61].

Although it is surprising that the *vgu* operon mutant is attenuated for virulence in the leaf, while being hyper-virulent in the tuber, the two *in planta* environments are very different. The tuber environment is anaerobic and contains a high level of starch, thus is a glucose rich environment, whereas the apoplastic spaces of the leaf are aerobic and nutrient-poor [67]. Demonstrably, nutrient availability and other environmental factors have dramatic effects on virulence expression. For instance, *hrpN* is regulated by the alternative sigma factor HrpL in addition to KdgR. HrpL is activated by a two-component system, HrpX/Y that is responsive to unknown exogenous metabolites. In tubers, HrpL is not up-regulated by

HrpX/Y and thus does not activate *hrpN*, which we see in our wild type expression.

CONCLUSION

The life history of *P. carotovorum*, spanning from a saprophytic environmental microbe to an opportunistic phytopathogen, suggests strong pressure to sense shifts in the environment and to respond accordingly. Despite the large amount of data on the network of regulators that tightly control expression of virulence factors in *P. carotovorum,* it is still unclear what and how different environmental signals are monitored to determine virulence network expression. Here, we identified a metabolic pathway necessary for gluconate metabolism that acts as a direct or indirect cue for virulence regulation.

LITERATURE CITED

1. McCarter-Zorner, N.J., Franc, G. D., Harrison, M. D., Michaud, J. E., Quinn, C. E., Sells, I. A., and Graham, D. C., *Soft rot Erwinia in surface and underground waters in Southern Scotland and in Colorado, United States.* J. Appl. Bacteriol., 1984. **57**: p. 95-105.

2. McCarter-Zorner, N.J., Harrison, M. D., Franc, G. D., Quinn, C. E., Sells, I. A., and Graham, D. C., *Soft rot Erwinia bacteria in the rhizosphere of weeds and crop plants in Colorado, United States and Scotland.* J. Appl. Bacteriol. , 1985. **59**: p. 357-368.

3. Harrison, M., Quinn, C., Ann Sells, I., and Graham, D., *Waste potato dumps as sources of insects contaminated with soft rot coliform bacteria in relation to re-contamination of pathogen-free potato stocks.* Potato Res., 1977. **20**(1): p. 37-52.

4. Molina, J.M., Harrison, D., and Brewer, J. W., *Transmission of Erwinia carotovora var. atroseptica by Drosophila melanogaster Meig. I. Acquisition and transmission of the bacterium.* Am. Potato J., 1974. **51**: p. 245-250.

5. Allen, C., S. Reverchon, and J. Robert-Baudouy, *Nucleotide sequence of the Erwinia chrysanthemi gene encoding 2-keto-3-deoxygluconate permease.* Gene, 1989. **83**(2): p. 233-41.

6. Matsumoto, H., et al., *Comparative study of regulatory mechanisms for pectinase production by Erwinia carotovora subsp. carotovora and Erwinia chrysanthemi.* Mol. Plant Microbe Interact., 2003. **16**(3): p. 226-37.

7. Lagaert, S., T. Belien, and G. Volckaert, *Plant cell walls: Protecting the barrier from degradation by microbial enzymes.* Semin. Cell Dev. Biol., 2009. **20**(9): p. 1064-73.

8. Murata, H., et al., *Regulation of the production of extracellular pectinase, cellulase, and protease in the soft rot bacterium Erwinia carotovora subsp. carotovora: evidence that aepH of E. carotovora subsp. carotovora 71 activates gene expression in E. carotovora subsp. carotovora, E. carotovora subsp. atroseptica, and Escherichia coli.* Appl Environ Microbiol, 1994. **60**(9): p. 3150-9.

9. Toth, I.K. and P.R. Birch, *Rotting softly and stealthily.* Curr. Opin. Plant Biol., 2005. **8**(4): p. 424-9.

10. Barnard, A.M. and G.P. Salmond, *Quorum sensing in Erwinia species.* Anal Bioanal Chem, 2007. **387**(2): p. 415-23.

11. Charkowski, A.O., *Decaying signals: will understanding bacterial-plant communications lead to control of soft rot?* Curr. Opin. Biotechnol., 2009. **20**(2): p. 178-84.

12. Flego, D., et al., *A two-component regulatory system, pehR-pehS, controls endopolygalacturonase production and virulence in the plant pathogen Erwinia carotovora subsp. carotovora.* Mol Plant Microbe Interact, 2000. **13**(4): p. 447-55.

13. Laasik, E., L. Andresen, and A. Mae, *Type II quorum sensing regulates virulence in Erwinia carotovora ssp. carotovora.* FEMS Microbiol Lett, 2006. **258**(2): p. 227-34.

14. Liu, H., et al., *Quorum sensing coordinates brute force and stealth modes of infection in the plant pathogen Pectobacterium atrosepticum.* PLoS Pathog., 2008. **4**(6): p. e1000093.

15. Pirhonen, M., et al., *A small diffusible signal molecule is responsible for the global control of virulence and exoenzyme production in the plant pathogen Erwinia carotovora.* Embo J, 1993. **12**(6): p. 2467-76.

16. Cui, Y., et al., *rsmC of the soft-rotting bacterium Erwinia carotovora subsp. carotovora negatively controls extracellular enzyme and harpin(Ecc) production and virulence by modulating levels of regulatory RNA (rsmB) and RNA-binding protein (RsmA).* J Bacteriol, 1999. **181**(19): p. 6042-52.

17. Liu, Y., et al., *kdgREcc negatively regulates genes for pectinases, cellulase, protease, HarpinEcc, and a global RNA regulator in Erwinia carotovora subsp. carotovora.* J Bacteriol, 1999. **181**(8): p. 2411-21.

18. Rodionov, D.A., M.S. Gelfand, and N. Hugouvieux-Cotte-Pattat, *Comparative genomics of the KdgR regulon in Erwinia chrysanthemi 3937 and other gamma-proteobacteria.* Microbiology, 2004. **150**(Pt 11): p. 3571-90.

19. Hugouvieux-Cotte-Pattat, N. and J. Robert-Baudouy, *Hexuronate catabolism in Erwinia chrysanthemi.* J. Bacteriol., 1987. **169**(3): p. 1223-31.

20. Nasser, W., J. Robert-Baudouy, and S. Reverchon, *Antagonistic effect of CRP and KdgR in the transcription control of the Erwinia chrysanthemi pectinolysis genes.* Mol. Microbiol., 1997. **26**(5): p. 1071-82.

21. Liu, Y., et al., *Characterization of a novel RNA regulator of Erwinia carotovora ssp. carotovora that controls production of extracellular*

enzymes and secondary metabolites. Mol Microbiol, 1998. **29**(1): p. 219-34.

22. Baker, C.S., et al., *CsrA regulates glycogen biosynthesis by preventing translation of glgC in Escherichia coli.* Mol Microbiol, 2002. **44**(6): p. 1599-610.

23. Mukherjee, A., et al., *RpoS (sigma-S) controls expression of rsmA, a global regulator of secondary metabolites, harpin, and extracellular proteins in Erwinia carotovora.* J Bacteriol, 1998. **180**(14): p. 3629-34.

24. Cui, Y., et al., *Erwinia carotovora subspecies produce duplicate variants of ExpR, LuxR homologs that activate rsmA transcription but differ in their interactions with N-acylhomoserine lactone signals.* J Bacteriol, 2006. **188**(13): p. 4715-26.

25. Cui, Y., et al., *ExpR, a LuxR homolog of Erwinia carotovora subsp. carotovora, activates transcription of rsmA, which specifies a global regulatory RNA-binding protein.* J Bacteriol, 2005. **187**(14): p. 4792-803.

26. Sjoblom, S., et al., *Cooperation of two distinct ExpR regulators controls quorum sensing specificity and virulence in the plant pathogen Erwinia carotovora.* Mol Microbiol, 2006. **60**(6): p. 1474-89.

27. Cui, Y., et al., *Identification of a global repressor gene, rsmA, of Erwinia carotovora subsp. carotovora that controls extracellular enzymes, N-(3-oxohexanoyl)-L-homoserine lactone, and pathogenicity in soft-rotting Erwinia spp.* J Bacteriol, 1995. **177**(17): p. 5108-15.

28. Chatterjee, A., et al., *Inactivation of rsmA leads to overproduction of extracellular pectinases, cellulases, and proteases in Erwinia carotovora subsp. carotovora in the absence of the starvation/cell density-sensing signal, N-(3-oxohexanoyl)-L-homoserine lactone.* Appl Environ Microbiol, 1995. **61**(5): p. 1959-67.

29. Lazdunski, A.M., I. Ventre, and J.N. Sturgis, *Regulatory circuits and communication in Gram-negative bacteria.* Nat Rev Microbiol, 2004. **2**(7): p. 581-92.

30. Whitehead, N.A., et al., *The regulation of virulence in phytopathogenic Erwinia species: quorum sensing, antibiotics and ecological considerations.* Antonie Van Leeuwenhoek, 2002. **81**(1-4): p. 223-31.

31. Yang, S., et al., *Type III secretion system genes of Dickeya dadantii 3937 are induced by plant phenolic acids.* PLoS One, 2008. **3**(8): p. e2973.

32. Li, Y., et al., *The plant phenolic compound p-coumaric acid represses gene expression in the Dickeya dadantii type III secretion system.* Appl. Environ. Microbiol., 2009. **75**(5): p. 1223-8.

33. Cui, Y., et al., *Regulatory network controlling extracellular proteins in Erwinia carotovora subsp. carotovora: FlhDC, the master regulator of flagellar genes, activates rsmB regulatory RNA production by affecting gacA and hexA (lrhA) expression.* J. Bacteriol., 2008. **190**(13): p. 4610-23.

34. Harris, S.J., et al., *The hexA gene of Erwinia carotovora encodes a LysR homologue and regulates motility and the expression of multiple virulence determinants.* Mol. Microbiol., 1998. **28**(4): p. 705-17.

35. Mukherjee, A., et al., *hexA of Erwinia carotovora ssp. carotovora strain Ecc71 negatively regulates production of RpoS and rsmB RNA, a global regulator of extracellular proteins, plant virulence and the quorum-sensing signal, N-(3-oxohexanoyl)-L-homoserine lactone.* Environ Microbiol, 2000. **2**(2): p. 203-15.

36. de Souza, J.T., M. Mazzola, and J.M. Raaijmakers, *Conservation of the response regulator gene gacA in Pseudomonas species.* Environ. Microbiol., 2003. **5**(12): p. 1328-40.

37. Kitten, T., et al., *A newly identified regulator is required for virulence and toxin production in Pseudomonas syringae.* Mol Microbiol, 1998. **28**(5): p. 917-29.

38. Chatterjee, A., et al., *GacA, the response regulator of a two-component system, acts as a master regulator in Pseudomonas syringae pv. tomato DC3000 by controlling regulatory RNA, transcriptional activators, and alternate sigma factors.* Mol Plant Microbe Interact, 2003. **16**(12): p. 1106-17.

39. Pruss, B.M., et al., *FlhD/FlhC is a regulator of anaerobic respiration and the Entner-Doudoroff pathway through induction of the methyl-accepting chemotaxis protein Aer.* J. Bacteriol., 2003. **185**(2): p. 534-43.

40. Leatham, M.P., et al., *Mouse intestine selects nonmotile flhDC mutants of Escherichia coli MG1655 with increased colonizing ability and better utilization of carbon sources.* Infect. Immun., 2005. **73**(12): p. 8039-49.

41. Pruss, B.M., et al., *A complex transcription network controls the early stages of biofilm development by Escherichia coli.* J. Bacteriol., 2006. **188**(11): p. 3731-9.

42. Boch, J., et al., *Identification of Pseudomonas syringae pv. tomato genes induced during infection of Arabidopsis thaliana.* Mol. Microbiol., 2002. **44**(1): p. 73-88.

43. Grant, S.R., et al., *Subterfuge and manipulation: type III effector proteins of phytopathogenic bacteria.* Annu Rev Microbiol, 2006. **60**: p. 425-49.

44. Miriagou, V., et al., *IS26-associated In4-type integrons forming multiresistance loci in enterobacterial plasmids.* Antimicrob. Agents Chemother., 2005. **49**(8): p. 3541-3.

45. Zienkiewicz, M., et al., *Mosaic structure of p1658/97, a 125-kilobase plasmid harboring an active amplicon with the extended-spectrum beta-lactamase gene blaSHV-5.* Antimicrob. Agents Chemother., 2007. **51**(4): p. 1164-71.

46. Glasner, J.D., et al., *Niche-specificity and the variable fraction of the Pectobacterium pan-genome.* Mol. Plant Microbe Interact., 2008. **21**(12): p. 1549-60.

47. Rudd, K.E., *Linkage map of Escherichia coli K-12, edition 10: the physical map.* Microbiol. Mol. Biol. Rev., 1998. **62**(3): p. 985-1019.

48. Truesdell, S.J., et al., *Pathways for metabolism of ketoaldonic acids in an Erwinia sp.* J. Bacteriol., 1991. **173**(21): p. 6651-6.

49. Chang, J.H., et al., *A high-throughput, near-saturating screen for type III effector genes from Pseudomonas syringae.* Proc. Natl. Acad. Sci. U.S.A., 2005. **102**(7): p. 2549-54.

50. Horton, R.M., et al., *Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension.* Gene, 1989. **77**(1): p. 61-8.

51. Datsenko, K.A. and B.L. Wanner, *One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products.* Proc. Natl. Acad. Sci. U.S.A., 2000. **97**(12): p. 6640-5.

52. Rashid, M.H. and A. Kornberg, *Inorganic polyphosphate is needed for swimming, swarming, and twitching motilities of Pseudomonas aeruginosa.* Proc. Natl. Acad. Sci. U.S.A., 2000. **97**(9): p. 4885-90.

53. Takle, G.W., I.K. Toth, and M.B. Brurberg, *Evaluation of reference genes for real-time RT-PCR expression studies in the plant pathogen Pectobacterium atrosepticum.* BMC Plant Biol., 2007. **7**: p. 50.

54. Koek, M.M., et al., *Microbial metabolomics with gas chromatography/mass spectrometry.* Anal. Chem., 2006. **78**(4): p. 1272-81.

55. Ruijter G. D. G., a.J.V., *Determination of intermediary metabolites in Aspergillus niger.* J. Microbiol. Met., 1996. **25**(3): p. 295-302.

56. Lindeberg, M., et al., *Proposed guidelines for a unified nomenclature and phylogenetic analysis of type III Hop effector proteins in the plant pathogen Pseudomonas syringae.* Mol. Plant Microbe Interact., 2005. **18**(4): p. 275-82.

57. Chatterjee, A., Y. Cui, and A.K. Chatterjee, *Regulation of Erwinia carotovora hrpL(Ecc) (sigma-L(Ecc)), which encodes an extracytoplasmic function subfamily of sigma factor required for expression of the HRP regulon.* Mol Plant Microbe Interact, 2002. **15**(9): p. 971-80.

58. Murray, E.L. and T. Conway, *Multiple regulators control expression of the Entner-Doudoroff aldolase (Eda) of Escherichia coli.* J. Bacteriol., 2005. **187**(3): p. 991-1000.

59. Eisenberg, R.C. and W.J. Dobrogosz, *Gluconate metabolism in Escherichia coli.* J. Bacteriol., 1967. **93**(3): p. 941-9.

60. Dong, X., *NPR1, all things considered.* Curr. Opin. Plant Biol., 2004. **7**(5): p. 547-52.

61. Barnard, A.M., et al., *Quorum sensing, virulence and secondary metabolite production in plant soft-rotting bacteria.* Philos. Trans. R. Soc. Lond. B Biol. Sci., 2007. **362**(1483): p. 1165-83.

62. Mole, B.M., et al., *Global virulence regulation networks in phytopathogenic bacteria.* Trends Microbiol., 2007. **15**(8): p. 363-71.

63. Pouyssegur, J. and F. Stoeber, *Genetic control of the 2-keto-3-deoxy-d-gluconate metabolism in Escherichia coli K-12: kdg regulon.* J. Bacteriol., 1974. **117**(2): p. 641-51.

64. Nasser, W., et al., *Inducing properties of analogs of 2-keto-3-deoxygluconate on the expression of pectinase genes of Erwinia chrysanthemi.* FEMS Microbiol. Lett., 1991. **65**(1): p. 73-8.

65. Smith, T.G. and T.R. Hoover, *Deciphering bacterial flagellar gene regulatory networks in the genomic era.* Adv. Appl. Microbiol., 2009. **67**: p. 257-95.

66. Chatterjee, A., Y. Cui, and A.K. Chatterjee, *RsmC of Erwinia carotovora subsp. carotovora negatively controls motility, extracellular protein production, and virulence by binding FlhD and modulating transcriptional activity of the master regulator, FlhDC.* J. Bacteriol., 2009. **191**(14): p. 4582-93.

67. Rico, A. and G.M. Preston, *Pseudomonas syringae pv. tomato DC3000 uses constitutive and apoplast-induced nutrient assimilation pathways to catabolize nutrients that are abundant in the tomato apoplast.* Mol. Plant Microbe Interact., 2008. **21**(2): p. 269-82.

Chapter 5: Discussion

An over-arching question in the study of *P. carotovorum* pathogenesis is how it transitions from an environmental microbe to a pathogen of various plant hosts and then coordinately deploys stealth virulence determinants, like the T3SS, and brute-force, like the T2SS [1]. In the chapters of this dissertation, I reviewed existing literature on the network of regulators that govern virulence in *P. carotovorum*, examined its transcriptional profiles during different infection routes, and identified a sugar metabolic pathway that acts as a differential environmental input to virulence responses. In this short discussion chapter I review the questions that have been answered, identify new and unanswered questions, and briefly posit the next hypotheses and some experimental designs that might be used to address them.

A central question addressed in the third chapter was what role the T3SS plays in the life-history of a macerating necrotroph such as *P. carotovorum*. In phytopathogens, T3SS generally function to manipulate host defenses so that the pathogen can grow unabated while keeping the host alive [2-5]. However, *P. carotovorum* kills its host with macerating enzymes [6], making the adaptive process of host-manipulation seem unnecessary. Using a combination of a promoter-trap screen [7] and a DNA microarray, we identified a small set of genes that are co-regulated with the T3SS in order to identify additional T3SS effectors. Two of the eight genes we found to be co-regulated with the T3SS may

encode candidate translocated effector proteins, one of which is a hypothetical protein and the other a putative transported protein. However, the other newly-identified HrpL regulon members do not look like likely candidates for T3SS effectors, for example genes predicted to encode prokaryotic transcription factors. Additionally, the transcriptional profiles of *P. carotovorum* during different infections indicate that the T3SS is only expressed and required for infection during leaf invasions. This agrees with my pathology data, which indicate that a T3SS deficient mutant has no attenuated virulence in tubers, but significant attenuation in leaf infections [8]. Lastly, data from our collaborators indicates that the known T3SS effector delivered by *P. carotovorum*, DspE, functions as a cytotoxin, does not alter host defenses, and is not recognized by R proteins in *Nicotiana benthamiana* (unpublished data). More specifically, the transcriptional profile of *N. bethamiana* infected with *P. carotovorum* matches that of *N. benthamiana* infected with a T3SS-deficient *P. syringae* strain, rather than a strain expressing a T3SS and delivering a suite of effectors that are suppressing defense responses.

From this data, we can now hypothesize that the T3SS in *P. carotovorum* is not required for suppression of basal defense responses, but has been adapted in this necrotrophic pathogen to simply and quickly deliver toxins to cells at the site of invasion during early stages of leaf infection. The leaf apoplast is a nutrient-poor environment [9], thus reaching a quorum for macerating enzyme production

and successful host invasion may be a slow process that requires the actions of a T3SS to essentially bide time.

This is not the case, perhaps, in the starch- (glucose) -rich tuber or the stem where *P. carotovorum* can invade the sucrose-rich phloem. Nutrient availability, more specifically, carbon sources, have long been established as key environmental signals for invasion, and particularly for deployment of T3SS [5, 10, 11]. However, *P carotovorum* may use carbon source availability to not only delineate host and non-host environments, but also to finely distinguish between infection routes and adjust infection strategy to deploy virulence determinants accordingly.

In the fourth chapter, I began to address this hypothesis by studying an operon in *P. carotovorum* that was previously associated with T3SS expression in *P. syringae* [12] and appeared to encode enzymes involved in sugar catabolism, as predicted by BLAST analysis. I found that the *vgu* operon is required for gluconate metabolism, and in its absence many key virulence regulators (introduced in chapter 1) were misregulated and virulence responses downstream were accordingly disrupted *in planta.* Additionally, our data are in agreement with previously reported data [13] that indicates that the gluconate metabolic pathway acts antagonistically to the pectin catabolic pathway, which is also known to be a signal for virulence responses, via KdgR activity [14-18]. These data support the idea that carbon sources are finely monitored upstream

of virulence regulation and that metabolic fluxes are regulated by intertwined catabolic pathways.

Of course, many questions are raised and remain unanswered by these data. Further work should investigate the function and potential virulence contribution of the HrpL-regulated genes identified in chapter 3. These novel T3SS-associated virulence determinants could help us better understand how this common virulence system has been adapted for the pathogenesis strategy of *P. carotovorum.* In particular, a few of the candidates encode proteins that have homology to transcription regulators, suggesting secondary levels of virulence regulation that have been integrated from this horizontally acquired system. Deletion mutations in these genes and subsequent pathology assays could assess their contribution to virulence regulation, transcriptomic studies could identify regulon members, and computational anylsis could be used to identify targeted *cis* elements.

The most striking unanswered questions from the discussion in chapter 3 is the identity of specific carbon sources monitored in *P. carotovorum* and how those signals are perceived and translated to inputs for the virulence regulation network. Correlating the transcriptional profile data given here from *in planta* conditions with transcriptome data under different carbon source availability conditions may be one way to identify other genes involved in metabolism and virulence signaling. However, a more direct experimental approach might be to

clarify metabolic pathways using isotope tracing, which involves feeding carbon isotope labeled substrates (examples might be galacturonate, the main component of pectin, or maltose, a disaccharide of glucose) to *P. carotovorum* and tracing the fate of the isotope with mass spectrometry or nuclear magnetic resonance (NMR) and inferring metabolic flux ratios [19]. Comparing the metabolic flux ratios of wild type *P. carotovorum* to that of the *vgu* operon mutant given a labeled substrate upstream of gluconate metabolism (which the mutant can metabolize, such as glucose) would clarify the role of the *vgu* operon in gluconate metabolism and identify coupled metabolic pathways. This would also identify catabolic intermediates that may act as direct signals for virulence regulators. Enzyme assays and functional analysis of key enzymes or candidate signal receptors by making point mutations of conserved domains could further refine our understanding of the proteins responsible for metabolism and virulence signaling.

In direct response to the over-arching question addressed in this dissertation, the data presented suggest that: carbon source discernment are key factors determining how *P. carotovorum* transitions from an environmental microbe to a pathogen of various infection routes; and that *P. carotovorum* coordinately deploys the T3SS only during the early stages of leaf infections, before resorting to brute-force virulence determinants, such as the T2SS.

LITERATURE CITED:

1. Toth, I.K. and P.R. Birch, *Rotting softly and stealthily.* Curr Opin Plant Biol, 2005. **8**(4): p. 424-9.

2. Grant, S.R., et al., *Subterfuge and manipulation: type III effector proteins of phytopathogenic bacteria.* Annu Rev Microbiol, 2006. **60**: p. 425-49.

3. Mudgett, M.B., *New insights to the function of phytopathogenic bacterial type III effectors in plants.* Annu Rev Plant Biol, 2005. **56**: p. 509-31.

4. Rahme, L.G., M.N. Mindrinos, and N.J. Panopoulos, *Plant and environmental sensory signals control the expression of hrp genes in Pseudomonas syringae pv. phaseolicola.* J Bacteriol, 1992. **174**(11): p. 3499-507.

5. Tang, X., Y. Xiao, and J.M. Zhou, *Regulation of the type III secretion system in phytopathogenic bacteria.* Mol Plant Microbe Interact, 2006. **19**(11): p. 1159-66.

6. Roberts, D.P., et al., *Requirement for two or more Erwinia carotovora subsp. carotovora pectolytic gene products for maceration of potato tuber tissue by Escherichia coli.* J Bacteriol, 1986. **167**(1): p. 279-84.

7. Chang, J.H., et al., *A high-throughput, near-saturating screen for type III effector genes from Pseudomonas syringae.* Proc Natl Acad Sci U S A, 2005. **102**(7): p. 2549-54.

8. Kim, H.S., et al., *Phylogeny and virulence of naturally occurring type III secretion system-deficient Pectobacterium strains.* Appl Environ Microbiol, 2009. **75**(13): p. 4539-49.

9. Rico, A. and G.M. Preston, *Pseudomonas syringae pv. tomato DC3000 uses constitutive and apoplast-induced nutrient assimilation pathways to catabolize nutrients that are abundant in the tomato apoplast.* Mol Plant Microbe Interact, 2008. **21**(2): p. 269-82.

10. Huynh, T.V., D. Dahlbeck, and B.J. Staskawicz, *Bacterial blight of soybean: regulation of a pathogen gene determining host cultivar specificity.* Science, 1989. **245**(4924): p. 1374-7.

11. Wei, Z., J.F. Kim, and S.V. Beer, *Regulation of hrp genes and type III protein secretion in Erwinia amylovora by HrpX/HrpY, a novel two-component system, and HrpS.* Mol Plant Microbe Interact, 2000. **13**(11): p. 1251-62.

12. Boch, J., et al., *Identification of Pseudomonas syringae pv. tomato genes induced during infection of Arabidopsis thaliana.* Mol. Microbiol., 2002. **44**(1): p. 73-88.

13. Nasser, W., et al., *Inducing properties of analogs of 2-keto-3-deoxygluconate on the expression of pectinase genes of Erwinia chrysanthemi.* FEMS Microbiol. Lett., 1991. **65**(1): p. 73-8.

14. Hugouvieux-Cotte-Pattat, N. and J. Robert-Baudouy, *Hexuronate catabolism in Erwinia chrysanthemi.* J. Bacteriol., 1987. **169**(3): p. 1223-31.

15. Liu, Y., et al., *$kdgR_{Ecc}$ negatively regulates genes for pectinases, cellulase, protease, $Harpin_{Ecc}$, and a global RNA regulator in Erwinia carotovora subsp. carotovora.* J. Bacteriol., 1999. **181**(8): p. 2411-21.

16. Nasser, W., J. Robert-Baudouy, and S. Reverchon, *Antagonistic effect of CRP and KdgR in the transcription control of the Erwinia chrysanthemi pectinolysis genes.* Mol. Microbiol., 1997. **26**(5): p. 1071-82.

17. Reverchon, S., W. Nasser, and J. Robert-Baudouy, *Characterization of kdgR, a gene of Erwinia chrysanthemi that regulates pectin degradation.* Mol. Microbiol., 1991. **5**(9): p. 2203-16.

18. Rodionov, D.A., M.S. Gelfand, and N. Hugouvieux-Cotte-Pattat, *Comparative genomics of the KdgR regulon in Erwinia chrysanthemi 3937 and other gamma-proteobacteria.* Microbiology, 2004. **150**(Pt 11): p. 3571-90.

19. Rantanen, A., et al., *An analytic and systematic framework for estimating metabolic flux ratios from 13C tracer experiments.* BMC Bioinformatics, 2008. **9**: p. 266.