AN ABSTRACT OF THE DISSERTATION OF

William J. Thomas for the degree of Doctor of Philosophy in Molecular and Cellular Biology presented on May 4, 2012.

Title: Identification and Characterization of Type III Effector Proteins in Plant-Associated Bacteria

Abstract approved:

_____

Jeff H. Chang

Symbioses between microbes and multicellular eukaryotes are found in all biomes, and encompass a spectrum of symbiotic lifestyles that includes parasitism and disease, commensalism, and mutually beneficial interdependent host-microbe relationships. Regardless of outcome, these symbiotic lifestyles are governed by a complex molecular "courtship" between microbe and potential host. This courtship is the primary determinant of the host range of a given microsymbiont. Host immunity poses a formidable barrier to the establishment of host-microbe relationships, and the majority of microbial suitors will be thwarted by it. Only by successfully "wooing" the host cell's immune defenses with the appropriate molecular signals can a microsymbiont successfully colonize its host.

A strategy common to microsymbionts across the spectrum of symbiotic lifestyles and host organisms is the delivery of microbial-encoded effector proteins into the cytoplasm of host cells to manipulate the host cell's molecular machinery for the purposes of subverting host immunity. Bacteria, in particular, have adapted a number of secretion systems for this purpose. The most well-characterized of these is the type III secretion system (T3SS), a molecular apparatus that specializes in injecting type III effector (T3Es) proteins directly into host cells. The work in this thesis focuses on T3Es of plant-associated bacteria, with particular emphasis on mutualistic bacteria. We present

evidence that collections of T3Es from *Sinorhizobium fredii* and *Bradyrhizobium japonicum* are, in stark contrast to those of phytopathogenic bacteria, in a co-evolutionary equilibrium with their hosts. This equilibrium is characterized by highly conserved T3E collections consisting of many "core" T3Es with little variation in nucleotide sequence. The T3Es of *Mesorhizobium loti* MAFF303099 suggest a completely different picture of the evolution of T3Es. MAFF303099 recently acquired its T3SS locus, and the work in this thesis provides an evolutionary snapshot of a mutualist that is innovating a T3E collection primarily through horizontal gene transfer. Collectively, this work represents the first comprehensive catalog of T3Es of rhizobia and, in the case of *Sinorhizobium* and *Bradyrhizobium,* the first evidence of purifying selection for T3Es.

Identification and Characterization of Type III Effector Proteins in Plant-Associated Bacteria

by

William J. Thomas

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of

the requirements for the

degree of

Doctor of Philosophy

Presented May 4, 2012

Commencement June 2012

Doctor of Philosophy dissertation of William J. Thomas presented on May 4, 2012.

APPROVED:

_____

Major Professor, representing Molecular and Cellular Biology

_____

Director of the Molecular and Cellular Biology Program

_____

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

_____

William J. Thomas, Author

ACKNOWLEDGEMENTS

I owe a tremendous debt of gratitude to the many people who have supported and guided me during the course of my graduate career. I would like to express my sincere appreciation to my adviser, Jeff Chang, for taking a chance on me and persevering with me, and for delivering criticism and accolades with equal enthusiasm. Jeff has demonstrated many times over that a mentor gives unstintingly of himself to his protégés. Likewise, to my friend, colleague, and collaborator Jeffrey Kimbrel, for his advice, help, and moral support, I offer my heartfelt appreciation. I would also like to extend my gratitude to Jason Cumbie for his coaching on all matters computational and for his collaboration on AutoSPOT, the program that made my life so much easier. To those who toiled beside me at the bench—Allison Creason, Andres Alvarez, Meesha Peña, Stanley Lee, David Swader, Ryan Lilley, Jayme Stout, and Gleb Basilewsky—many thanks for your hard work, which helped to erect the foundation upon which much of this thesis is built. Thanks also to the members of my graduate committee—Joyce Loper, Kerry McPhail, Martin Schuster, Virginia Weis, and Tom Wolpert—who made each stage of my education a positive learning experience. I owe a special thanks to my old mentor, Dr. Chester Leathers, for his confidence in me and his encouragement, which helped me to embark on this journey. Most importantly, I wish to thank my family: my parents, brother, aunts, uncle, and in-laws, for every conceivable type of support; my children, Owen and Lucy, for keeping my future bright; and my amazing wife, Heidi, who makes it all possible.

CONTRIBUTION OF AUTHORS

Chapter 1: William J. Thomas wrote the chapter.

Chapter 2: William J. Thomas and Jeffrey A. Kimbrel prepared the DNA for sequencing and assembled the genomes. William J. Thomas, with assistance from Jeffrey A. Kimbrel, Allison L. Creason, and Caitlin A. Thireault, cloned candidate effector proteins and assayed them for translocation. Jeffrey A. Kimbrel conducted subsequent bioinformatics analyses. William J. Thomas, Jeffrey A. Kimbrel, Jeff H. Chang, and Joel L. Sachs wrote the manuscript.

Chapter 3: Ryan R. Lilley cloned candidate effector proteins. William J. Thomas assayed the candidate effector clones for translocation. William J. Thomas and Jeffrey A. Kimbrel conducted subsequent bioinformatics analyses. William J. Thomas, Jeffrey A. Kimbrel, and Jeff H. Chang wrote the manuscript.

Chapter 4: Caitlin A. Thireault replaced the tetracycline resistance gene from the mini-Tn5 vector with a kanamycin resistance gene flanked by FLP recognition target sites and integrated the construct into *Pseudomonas fluorescens* Pf0-1 (designated EtHAn). Caitlin A. Thireault, with assistance from William J. Thomas, cloned 30 individual effector proteins from *P. syringae* pv. *tomato* DC3000 into EtHAn. William J. Thomas did subsequent gene expression analyses of effector proteins in EtHAn, as well as assays for delivery of effector proteins from EtHAn, effector-mediated changes in virulence of EtHAn, and perturbations of plant immunity readouts by EtHAn-delivered effector proteins. William J. Thomas and Jeff H. Chang wrote the manuscript.

Chapter 5: William J. Thomas wrote the chapter.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

LIST OF TABLES

Table Page

LIST OF APPENDIX FIGURES

Figure Page

Introduction: Plant-microbe interactions in symbiosis

INTRODUCTION

Colonization by microbial symbionts is ubiquitous among multicellular organisms. These symbioses pervade every habitat and likely all branches of the phylogenetic tree of life. The term "symbiosis" encompasses a variety of lifestyles resulting from specific host-microbe interactions with a spectrum of possible outcomes ranging from mutually beneficial interactions to pathogenic interactions. Regardless of the outcome of symbiosis, successful colonization of a host organism by a microbe is typically governed by a complex molecular "courtship", during which the microbe navigates host barriers to successfully establish a symbiotic relationship. Such courtships are rarely successful, and many microbial suitors are potentially thwarted by the defenses of potential hosts; nevertheless, the infrequent successes are sufficient to account for the dizzying diversity of host-microbe relationships observed in nature. This thesis examines the co-evolution and role of bacterial-encoded type III effector proteins in host range determination in the context of symbiosis with plants.

PLANT BARRIERS TO SYMBIOSIS

Plant immunity presents an imposing obstacle for potential symbionts. While plants lack the adaptive immunity found in mammals, they do share with other multicellular eukaryotes an innate immune system that is sufficient to resist colonization by the majority of microbes (Nürnbeger *et al.,* 2004; Postel and Kemmerling, 2009). In plants, this basal defense is mediated by the host cell's pattern recognition receptors (PRRs), which recognize conserved molecular patterns of microbes (Sanabria *et al.*, 2010; Segonzac and Zipfel, 2011). Historically, these conserved molecular patterns have been referred to as pathogen-associated molecular patterns, or PAMPs, reflecting the conventional view of plant immunity as a barrier to pathogen invasion. It has been proposed that these conserved molecular patterns should receive the more general appellation of microbe-associated molecular patterns, or MAMPs (Ausubel, 2005). This move to a more inclusive nomenclature is indicative of a shift in the understanding of plant immunity as a barrier against microbes irrespective of their symbiotic lifestyles.

Plant PRRs are plasma membrane-associated receptors that decorate the surface of the host cell (Sanabria *et al.*, 2010). Interaction between the extracellular domain of a PRR and its cognate ligand initiates a series of responses collectively referred to as PAMP- or MAMP-triggered immunity (PTI or MTI, respectively; Jones and Dangl, 2006; Schwessinger and Zipfel, 2008). The pathogen-derived signal is internalized and transduced through a MAP kinase cascade, which results in rapid transcriptional changes, including induction of pathogenesis-related (PR) genes (Schwessinger and Zipfel, 2008; Pitzschke *et al.*, 2009; Tena *et al.,* 2011). MTI also results in an immediate influx of Ca^{2+} into the cytoplasm of the plant cell. This cytosolic Ca^{2+} signal is transduced through multiple signaling pathways, resulting in a rapid burst of reactive oxygen species via activation of NADPH oxidase, calmodulin-mediated production of nitric oxide (NO), and a calcium-dependent protein kinase (CDPK) signaling cascade (Schwessinger and Zipfel, 2008; Segonzac and Zipfel, 2011; Ma, 2011; Tena *et al.,* 2011). These immediate responses further transduce and amplify pathogen-derived signals into other early-response pathways. For instance, the MAP kinase cascade activates production of the phytohormone ethylene, while CDPK activity is responsible for the activation of salicylic acid (SA), another phytohormone (Pitzschke *et al.*, 2009; Segonzac and Zipfel, 2011; Tena *et al.,* 2011). The phytohormones ethylene and SA, as well as jasmonic acid (JA), control major regulatory networks that both complement and antagonize each other (An and Mou, 2011; Robert-Seilaniantz *et al.*, 2011). The complex crosstalk between these regulatory networks coordinates the timing and amplitude of the physiological responses involved in plant immunity (Robert-Seilaniantz *et al.*, 2011).

Many of the regulatory pathways described above converge on cell wall defenses, a set of physiological responses that manifest as strengthening of the physical barrier between plant cell and microbes. The ROS produced during the rapid oxidative burst in the early stages of MTI directly strengthen host cell walls by cross-linking glycoproteins (Torres *et al.*, 2006). ROS signaling is also implicated in the deposition of callose to form cell wall-reinforcing papillae. The ROS signals initiating the callose response are themselves initiated via multiple pathways, including Ca^{2+}-activated NADPH

oxidase production and abscisic acid (ABA), another phytohormone with a role in MTI (Luna *et al.*, 2011; Cao *et al.*, 2011). Interestingly, ABA is also a negative regulator of the callose response in its role as an antagonist of SA-mediated callose deposition (Cao *et al.*, 2011). Because the callose response is downstream of many earlier MTI responses, it is a convenient readout for perturbations of host cell defenses.

MTI is a robust first line of defense against microbes, and is sufficient to establish resistance against microbial symbioses. Symbionts must dampen or evade this response in order to successfully establish a host-microbe relationship. A common strategy for suppression of MTI involves the deployment of effector proteins into the cytoplasm of the plant cell for the purpose of manipulating the cellular machinery therein (Chisholm *et al.*, 2006; Jones and Dangl, 2006). Bacteria, for example, can use several different secretion systems for the delivery of effector proteins into host cells. In particular, the type III secretion system (T3SS) has been intensively characterized in the context of its role in symbioses of both plants and animals. Each is used by plant-associated bacteria to deliver its respective effector proteins for the purposes of dampening MTI (Lewis *et al*, 2009; Cascales and Christie, 2003).

THE TYPE III SECRETION SYSTEM

The T3SS is widely distributed among Gram-negative plant-associated bacteria, including many agriculturally and economically important phytopathogens, as well as the equally-important group of mutualistic nitrogen-fixing bacteria classified as rhizobia (Grant *et al.*, 2006). The structure of the T3SS consists of a flagellar-type ring structure spanning the inner and outer membrane of the bacteria, and a pilus that forms a conduit from the ring structure to the membrane of the host cell (He *et al.*, 2004; Galán and Wolf-Watz, 2006). Type III effector proteins (T3Es) are translocated through this molecular conduit into the host cell (He *et al.*, 2004; Galán and Wolf-Watz, 2006). In the case of phytopathogens, it is well established that the translocated T3Es collectively manipulate the host's cellular machinery to dampen defense (Cunnac, 2009; Lewis, 2009). The combined perturbations of plant defense by a collection of T3Es reduce MTI

below the level required for effective resistance (Jones and Dangl, 2006). In mutualistic rhizobia, a handful of T3Es have been identified with a demonstrable role in suppression of MTI. NopE1 and NopE2 of the rhizobia species *Bradyrhizobium japonicum* and NopL of the rhizobia strain *Sinorhizobium fredii* NGR234 are required for full nodulation of their respective legume hosts (Wenzel *et al.*, 2010; Marie *et al.*, 2003). Furthermore, NopL has been shown to suppress nodule senescence in nodules on *Phaseolus vulgaris,* while its expression in *Nicotiana tabacum* interferes with MAPK signaling (Zhang *et al.*, 2011). Clearly, the role of T3Es in evading MTI-mediated resistance to infection is common across the full spectrum of symbiotic lifestyles.

RHIZOBIA: HOST-SPECIFIC SYMBIOSES

The rhizobia are a diverse collection of alpha-proteobacteria comprising over 50 species in seven genera of the family *Rhizobiaceae* (Willems, 2006). Each species is further divided into strains, many of which are able to form nitrogen-fixing nodules in compatible host genotypes. This symbiotic relationship begins with a series of complex molecule exchanges between the rhizobia and host plant roots and culminates in the formation of root nodules. Within these nodules, the bacteria persist in a differentiated form capable of reducing atmospheric nitrogen to ammonia (Spaink, 2000; Kereszt *et al*., 2011). This nitrogen fixation represents a critical component of the nitrogen cycle, establishing a crucial role for the rhizobia-legume relationship in both agricultural and natural ecosystems.

Multiple levels of specificity dictate the compatibility of the host-microbe interaction between plant and rhizobia. These levels of specificity mediate the possible outcomes of the interaction, only one of which is the formation of productive, nitrogen-fixing nodules (van Rhijn and Vanderleyden, 1995). The limited host range of rhizobia constitutes the first level of specificity. Legumes produce genotype-specific combinations of flavonoids that function as signaling molecules in the rhizosphere. Detection of flavonoids produced by a compatible legume host induces the expression of nodulation (*nod*) genes in rhizobia; conversely, flavonoids from non-host legumes may

suppress rhizobial *nod* gene expression (Perret *et al.*, 2000; Hirsch, 2001). Many flavonoid-induced *nod* genes synthesize Nod factors (NF), the lipo-chitooligosaccharides secreted by rhizobia during the initial stages of the plant-rhizobia interaction. Rhizobial NF are perceived by cognate receptors of compatible legume hosts, resulting in induction of host responses (calcium spiking, root-hair curling) required for infection thread formation (Gage and Margolin, 2000; Spaink, 2000; Jones *et al.*, 2007; Haney *et al.*, 2011)

The lipopolysaccharides (LPSs) and exopolysaccharides (EPSs) that decorate the rhizobial outer membrane are also hypothesized to function in host specificity (Brzoska and Signer, 1991; Dazzo *et al.*, 1991; López-Lara *et al.*, 1995; Jones *et al.*, 2007; Hidalgo *et al.*, 2010). Legumes encode a distinct family of lectins with conserved carbohydrate recognition domains, which may recognize compatible rhizobial LPS and EPS (De Hoff *et al.*, 2009). Interestingly, lectins are also hypothesized to function in plant immunity, indicating a potential dual role in which lectins induce symbiosis responses or trigger defenses based on recognition of rhizobial LPS or EPS (De Hoff *et al.*, 2009).

Additionally, T3SS, which affect host range of pathogens, are demonstrably used by many strains of rhizobia, including all known strains of *B. japonicum*, *S. fredii* strains NGR234, USDA207, USDA257, and USDA191, and *Mesorhizobium loti* MAFF303099 (Bellato *et al.*, 1997; Freiberg *et al.*, 1997; Viprey *et al.*, 1998; Kaneko *et al.*, 2000). In these rhizobia, T3SS-associated genes are likely expressed early during symbiosis since they are co-regulated with the *nod* genes. Flavonoids are perceived in rhizobia cells as aglycones, which bind to the LysR-type transcriptional regulator NodD1, which in turn activates expression of both *nod* genes and T3SS-associated genes (Viprey *et al.*, 1998; Hirsch, 2001; Krause *et al.*, 2002; Wassem *et al.*, 2008). Given the demonstrable role of type III effectors during pathogenesis, a parallel role has been hypothesized for mutualism with T3SS-related proteins and host plant defenses forming another level of host specificity (Fauvart and Michiels, 2008, Deakin and Broughton, 2009; Soto *et al.*, 2009; Zamioudis and Pieterse, 2011).

TYPE III EFFECTOR PROTEINS AND HOST RANGE DETERMINATION

Individual T3Es can trigger another level of plant immunity. Plants have disease resistance proteins (R proteins) that perceive microbial effector proteins. Most *R* genes, such as *RPS2*, encode for proteins with nucleotide binding (NB) and leucine-rich repeat (LRR) motifs (Joshi and Nayak, 2011). In incompatible interactions, a specific T3E is recognized by a plant cell-encoded cognate R protein, inducing effector-triggered immunity (ETI, aka "gene-for-gene resistance" Jones and Dangl, 2006; Zhou, 2008; Mansfield, 2009; Rafiqi *et al.*, 2009). The set of responses induced by the effector-derived signal resembles MTI, but differs in the increased amplitude and accelerated timing of the induction (Jones and Dangl, 2006). Additionally, ETI is often associated with a hypersensitive response (HR) that is characterized by programmed cell death (Greeenberg and Yao, 2004). A T3E that elicits ETI on an incompatible host is referred to as an "avirulence factor" because of its role as a negative determinant of host range. Indeed, many T3E genes were given *avr* designations based on this discovery before they were shown to function in virulence on compatible hosts (Cui *et al.*, 2009).

The "gene-for-gene resistance" phenomenon has primarily been studied in the context of pathogenic symbioses, but has also been suggested in legume-rhizobia interactions. A role for gene-for-gene resistance in host-range specificity of rhizobia is supported by the identification of rhizobial T3Es with potential avirulence protein-like behaviors. NopP, which is delivered by *S. fredii* NGR234 into host plant cells, is required for full nodulation of *Pachyrhizus tuberosus, Tephrosia vogelii,* and *Flemingia congesta*; however, *nopP* mutants of NGR234 demonstrate increased nodulation of *Vigna unguiculata*, an incompatible host of wild type NGR234 (Schechter *et al.*, 2010; Ausmees *et al.*, 2004; Skorpil *et al.*, 2005). Similarly, translocation of the *B. japonicum* T3Es NopE1 and NopE2 is necessary for nodulation of its host, *Macroptilium atropurpureum,* yet has been shown to restrict nodulation in non-host *Vigna radiata* (Wenzel *et al.*, 2010). The T3SS-secreted proteins NopT and Y4lO of NGR234 are required for full nodulation of *Crotalaria juncea* and *V. unguiculata*, respectively, yet mutants affected in either of their corresonding genes show increased nodulation in *Phaseolus vulgaris* and *T. vogelii* (Dai *et*

al., 2008; Yang *et al.*, 2009). Furthermore, despite differing host ranges in the *S. fredii* strains USDA257, USDA191, and HH103, loss of the T3SS-secreted protein NopX results in compatibility with *Erythina* (Bellato *et al.*, 1997). In soybean, two *R* genes, *Rfg1* and *Rj2*, have been identified as loci that restrict nodulation by the rhizobial strains *S. fredii* USDA257 and *B. japonicum* USDA122, respectively (Yang *et al.*, 2010). The identification of nucleotide-binding leucine-rich repeat class provides tantalizing molecular data that these are indeed R proteins and support a gene-for-gene interaction with a recognized rhizobial effector as the basis for restriction of nodulation (Yang *et al.*, 2010).

CO-EVOLUTION OF T3ES AND PLANT DEFENSE

For plant pathogens, the co-evolution of host defense and T3Es has been modeled as an "evolutionary arms race" (Stavrinides *et al*., 2008). This model predicts rapid evolution of T3E collections in response to the selective pressures imposed by the host. The genetic patterns of T3E genes provide strong support for this model (Ma *et al.*, 2006; Zhou *et al.,* 2009; Baltrus *et al.*, 2011; Jackson *et al.*, 2011). A large-scale, genome-enabled analysis of 19 strains of the plant pathogen *Pseudomonas syringae* demonstrated considerable variation in the composition of T3E repertoires within this species, with little overlap between strains (Baltrus et al., 2011). Considerable sequence variation was also observed within the few conserved or core T3E families, providing further evidence of rapid evolution compatible with the arms race model (Baltrus *et al*., 2011). Additionally, collections of T3Es are redundant, which makes individual T3Es dispensable, thereby allowing phytopathogenic bacteria to respond to the selective pressures of *R* gene-mediated resistance by dispensing with perceived T3Es (Stavrinides *et al.*, 2008, Cunnac *et al.*, 2009; Cunnac *et al.*, 2011). Thus, the loss and gain of genes contribute to the dynamic nature of pathogen T3E collections.

While the evolutionary arms race between phytopathogenic bacteria and their plant hosts has considerable empirical support, a model for the evolution of T3Es in mutualistic bacteria has yet to emerge. One possibility is that the evolutionary arms race model will describe the legume-rhizobia symbiosis. In this scenario, plant and

microsymbiont compete with each other to increase the benefits obtained from the mutualism while minimizing the costs of participating in the relationship (Herre *et al.*, 1999; Douglas, 2008; Sachs *et al.*, 2010). An alternative model predicts an "evolutionary armistice" in which the partners of the legume-rhizobia partnership are under a purifying selection to maintain a co-evolutionary equilibrium (Sachs *et al.*, 2011a).

These alternative models are informed by two contrasting narratives observed in rhizobial evolution of T3SS loci. In *B. japonicum*, which nodulates soybean, the T3SS locus has been identified in all examined strains (Mazurier *et al.*, 2006). The degree of conservation of the T3SS locus within the species *B. japonicum* can be inferred to indicate an ancient acquisition of this mechanism that predates the divergence of the various strains (Mazurier *et al.*, 2006). By contrast, the lotus symbiont *M. loti* MAFF303099 is the only strain of *Mesorhizobium* in which a T3SS locus has been identified (Hubber *et al.*, 2004). Comparative sequence analysis of *M. loti* MAFF303099 to *M. loti* R7A indicates that, while most or all of the genes involved in nodulation and nitrogen fixation are shared by both strains, they differ significantly in mechanisms likely important for initiating the symbiotic relationship, inasmuch as R7A encodes a T4SS rather than a T3SS (Sullivan *et al.*, 2002; Uchiumi *et al.*, 2004). Furthermore, R7A is representative of other *M. loti* strains, all of which, with the exception of MAFF303099, encode a T4SS. The fact that MAFF303099 encodes a remnant of a T4SS-related gene just upstream of the T3SS locus indicates that the T4SS was present in this strain but has been lost, possibly in the same recombination event responsible for the acquisition of the T3SS in MAFF303099. These two narratives—the ancient lineage of the T3SS in *B. japonicum* and its recent acquisition in *M. loti* MAFF303099—have the potential to illuminate the selective forces driving the evolution of T3Es of rhizobia. Sequencing and comparison of genomes within this diverse phylogenetic framework is an important first step in elucidating evolutionary trends characterizing these selective pressures and corresponding changes in host range.

CONCLUSIONS

Chapters 2 and 3 of this thesis describe my work towards elucidating the evolutionary forces shaping the host range of rhizobia. The research presented therein represents the first and most extensive effort to catalog the T3Es of mutualist plant-associated bacteria. In Chapter 2, I describe this cataloging of T3Es for three strains of *S. fredii*, chosen for their broad host range and, in the case of strains USDA207 and USDA257, their T3SS-dependent host-range polymorphisms. In addition, we identified the T3Es of five diverse strains of *B. japonicum*, a species in which all analyzed strains encode a T3SS, making this species an ideal candidate for comparative genomics between strains. Interestingly, the findings described in Chapter 2 indicate that collections of T3Es of *S. fredii* and *B. japonicum* are highly conserved in both content and sequence. In this case, the paradigm for the co-evolution of plant host and microsymbiont closely corresponds to the model of an evolutionary armistice. In contrast, the findings presented in Chapter 3 describe the evolutionary forces shaping the innovation of a new T3E repertoire in the context of a relatively recently-acquired T3SS. The collections of T3Es identified using this combination of computational and functional approaches provide a starting point for understanding the role of the T3SS in host range determination, as well as a blueprint for identifying T3E repertoires in other plant mutualists.

Chapter 4 describes the cloning of an exogenous T3SS locus into a non-host-associated bacterium, *P. fluorescens*. The purpose of this construction was to develop a system for the heterologous delivery of T3Es in isolation for the purposes of characterizing their individual contributions to host-range determination and perturbations of host cell defenses. Using T3Es cloned from the phytopathogen *P. syringae* pv. *tomato* DC3000, we demonstrated the efficacy of this heterologous delivery system as a means of studying T3E-elicited phenotypes in plant hosts. In Appendix 2, I provide preliminary results that detail my first steps towards characterizing the sufficiency in a collection of T3Es for changing the host range of a pathogen, and for

converting the non-pathogenic EtHAn delivery system, described in Chapter 4, into a phytopathogenic bacterium.

In Chapter 5, I discuss the findings presented in this thesis and their implications for future research.

Extreme Evolutionary Stasis of Type III Effector Genes in Mutualistic *Sinorhizobium fredii* and *Bradyrhizobium japonicum*

Jeffrey A. Kimbrel*, William J. Thomas*, Allison L. Creason, Caitlin A. Thireault, Joel L. Sachs, Jeff H. Chang

*These authors contributed equally to the work

Manuscript in preparation for *PLoS Biology*

ABSTRACT

The co-evolution of mutualists and their eukaryotic hosts has been modeled according to two opposing paradigms. In one, mutualist and host are under conflict, with both maximizing fitness at the expense of the other. In the other, mutualist and host have reached resolution in their conflicts and benefit by enhancing the fitness of each other. We coupled the mining of eight finished and draft genome sequences to a functional screen to present a comprehensive cataloging of type III effectors and an empirical testing of the two opposing co-evolutionary frameworks in two species of mutualistic rhizobial species, *Sinorhizobium fredii* and *Bradyrhizobium japonicum*. Type III effectors are best known as virulence factors that are essential for Gram-negative pathogens to infect hosts and typically present evidence for dynamic evolution. Our characterizations of the confirmed type III effector genes show that rhizobial collections, in contrast to those of phytopathogens, are static, with strong conservation in content, little evidence for type III effector gain or loss, and little sequence divergence. In all, these data are consistent with a framework indicative of a mutualistic environment where microbe and host populations have stabilized.

AUTHOR SUMMARY

Rhizobia are an important group of bacteria that can enter into mutually beneficial symbiotic interactions with legume plants to fix atmospheric nitrogen. However, in order to do so, a complex dialog involving the exchange of chemical and molecular signals must occur between partners. Interestingly, some species of beneficial rhizobia employ a type III secretion system, a well-characterized virulence mechanism used by pathogens to inject bacterial-encoded type III effector proteins into hosts to manipulate host cells and dampen immunity. In this study, we generated draft genome sequences and employed computational as well as experimental methods to identify type III effectors from eight strains representing two species of rhizobia. We demonstrate that rhizobia have large collections of type III effectors, yet these collections are highly conserved in content with little diversity between strains. This

work represents the first investigation, to our knowledge, of type III effector content and evolution in the context of mutualism and is an important step towards understanding the roles for type III secretion systems and their effectors in mutualistic interactions.

INTRODUCTION

Eukaryotes universally encounter bacteria that inhabit, infect, and often provide significant benefits to host fitness (Medina and Sachs, 2001). In many cases, bacterial mutualist lineages likely share a deep evolutionary history with the hosts they infect, giving each partner opportunity to intimately shape the other's phenotype (Currie *et al.,* 2006; Brock *et al.,* 2011; Sachs *et al.*, 2011c). Two competing paradigms remain unresolved about the co-evolution of bacterial mutualists with their eukaryotic hosts (Sachs *et al*, 2011). One paradigm models host-mutualist interactions as an antagonistic arms race, as is the case for host-pathogen co-evolution. In this regard, even though both host and bacteria can attain net fitness benefits from host-mutualist interactions, natural selection is predicted to be relentless as it continuously shapes each partner to selfishly maximize their own gains from the interaction and minimize costs invoked by the other (Herre *et al.*, 1999; Douglas, 2008; Sachs *et al.*, 2010). The alternative framework predicts beneficial co-adaptation, with each partner population stabilized to an evolutionary détente (Parker, 1999; Law and Lewis, 2008; Sachs *et al.*, 2011c). Empirically testing these competing frameworks will help to resolve whether eukaryotic-bacterial mutualisms represent reciprocally exploitative interactions or if conflicts between bacterial mutualists and hosts have mostly been resolved (Herre *et al.*, 1999; Sachs *et al.*, 2004; Sachs and Simms, 2008).

A striking and well-studied example of arms race evolution occurs between Proteobacterial pathogens and plant hosts. Plants have multiple defense systems to recognize and respond to bacterial infection. One key plant defense is pattern-triggered immunity (PTI), in which pattern recognition receptors detect conserved microbe-associated molecular patterns and trigger a battery of adaptive responses (Jones and Dangl, 2006; Schwessinger and Zipfel, 2008; Segonzac and Zipfel, 2011). To counteract

host defenses, many Proteobacteria use type III secretion systems (T3SS) to deliver collections of bacterial-encoded type III effector proteins (T3Es) directly into host cells. For phytopathogens, evidence strongly indicates that T3Es are deployed to dampen host defenses, thereby allowing the bacteria to proliferate within host tissues and cause disease (Jones and Dangl, 2006; Cunnac *et al.*, 2009; Lewis *et al.*, 2009; Cui *et al.*, 2009; Mansfield, 2009; Rafiqi *et al.*, 2009). A second defense of plants is effector-triggered immunity (ETI) in which resistance (R) proteins perceive the presence or action of corresponding microbial effectors (Jones and Dangl, 2006; Cui *et al.*, 2009; Mansfield, 2009; Rafiqi *et al.*, 2009). The outputs of ETI and PTI are similar, but the former is associated with more robust and amplified responses and often a localized programmed cell death called the hypersensitive response (HR; Greenberg and Yao, 2004). Most *R* genes encode for nucleotide binding (NB) and leucine-rich repeat (LRR)-containing proteins. The perceived effectors were classically referred to as "avirulence proteins" (Avrs) because they betrayed the microbe to the plant.

Under the scenario of an antagonistic arms race, collections of bacterial T3E genes are predicted to exhibit patterns of genetic variation that reflect rapid evolution as hosts evolve defenses to the most commonly encountered strains. Indeed, analyses of plant pathogens have provided ample support for this model (Ma *et al.*, 2006; Zhou *et al.*, 2009; Baltrus *et al.*, 2011; Jackson *et al.*, 2011). In *Pseudomonas syringae*, for example, collections of T3E genes vary dramatically in size and content with few T3Es that would be considered "core" (Baltrus *et al.*, 2011). Another aspect of pathogen T3E collections is that their robustness is ensured via redundancy so that any individual T3E gene is dispensable (Cunnac *et al.*, 2009; Stavrinides *et al.*, 2008; Cunnac *et al.*, 2011). Hence, under the arms race scenario, rapid evolution of T3Es is advantageous to phytopathogens as novel collections of T3Es are more likely to avoid recognition while maintaining sufficiency in subverting host defenses.

Functional T3SS orthologs have also been uncovered in diverse species of Proteobacterial commensals and mutualists, including *P. fluorescens, Sodalis glossinidius, Sitophilus zeamais*, and numerous rhizobia, including *Sinorhizobium fredii* strains NGR234

(aka *Rhizobium* sp. NGR234), USDA191, USDA207, and USDA257, and *Bradyrhizobium japonicum*, as well as *Mesorhizobium loti* MAFF303099 (Bellato *et al.*, 1996; Freiberg *et al.*, 1997; Viprey *et al.*, 1998; Kaneko *et al.*, 2000; Dale *et al.*, 2001; Mazurier *et al.*, 2006; Rainey, 1999; Shigenobou *et al.*, 2000; Preston *et al.*, 2001; Dale *et al.*, 2002; Moran *et al.*, 2005; Kimbrel *et al.*, 2010). Many rhizobia are able to form nitrogen-fixing nodules in compatible host genotype lineages. T3SSs are known to be important for the establishment of bacterial mutualist infections even if the precise role of the T3SS is unknown. For instance, rhizobial mutants compromised in the construction or expression of the T3SS are often less efficient in establishing interactions with their otherwise compatible hosts (Meinhardt *et al.*, 1993; de Lyra *et al.*, 2006; Deng *et al.*, 2009). Moreover, loss-of-function mutants of rhizobial T3SS-secreted proteins and potential T3Es, such as NopT and NopJ (y4lO) of NGR234, are significantly compromised in nodulating their legume hosts (Annapurna and Krishnan, 2003; Marie *et al.*, 2003; Krishnan *et al.*, 2003; Ausmees *et al.*, 2004; Lorio *et al.*, 2004; Saad *et al.,* 2005; Rodrigues *et al.*, 2007; Dai *et al.*, 2008; López-Baena *et al.*, 2008; Zehner *et al.*, 2008; Hempel *et al.*, 2009; Kambara *et al.*, 2009; Yang *et al.*, 2009; Schechter *et al.*, 2010; Wenzel *et al.*, 2010).

Characterizations of rhizobial T3SSs have uncovered many parallels with T3SSs of bacterial pathogens. Rhizobial T3SS genes are co-regulated with the *nod* genes, pointing to an early role in host infection, as is the case with plant pathogenic bacteria (Viprey *et al.,* 1998; Krause *et al.*, 2002; Wassem *et al.*, 2008; Haapalainen *et al.*, 2009). The T3Es of rhizobia also affect bacterial host range, as evidenced by the repeated observations that rhizobia mutants deleted of confirmed or putative T3E genes have expanded host ranges beyond that of their corresponding wild type strain (Bellato *et al.*, 1997; Ausmees *et al.*, 2004; Skorpil *et al.*, 2005; Dai *et al.*, 2008; Yang *et al.*, 2009; Schechter *et al.*, 2010; Wenzel *et al.*, 2010). Finally, legume host loci responsible for "nodulation restriction" of rhizobia encode for NB-LRR proteins and map to regions with clusters of other NB-LRR-encoding genes linked to resistance against phytopathogens (Triplett and Sadowsky, 1992; Kanazin *et al.*, 1996; Graham *et al.*, 2002; Graham *et al.* 2002b; Yang *et al.*, 2010).

Here, we use whole-genome sequence data to investigate the molecular evolution of T3Es in mutualistic rhizobia and test the competing arms race versus evolutionary détente paradigms. We cataloged T3Es from *Sinorhizobium fredii* NGR234, USDA207, and USDA257, and *Bradyrhizobium japonicum* USDA6, USDA110, USDA122, USDA123, and USDA124. We chose these strains based on phylogenetic divergence and demonstrable reliance on T3SS for host infection. At the onset of our work, the only available finished genome sequence was from *S. fredii* NGR234 and *B. japonicum* USDA110, so we therefore used Illumina sequencing to generate draft genome sequences for the remaining strains (Freiberg *et al.*, 1997; Schmeisser *et al.*, 2009; Kaneko *et al.*, 2002). The draft and finished genome sequences were mined and candidate T3Es were confirmed for direct delivery into plant cells using a heterologous system. We used the confirmed T3Es to test the null hypothesis that collections of T3Es of mutualists evolve in a manner similar to collections of T3Es in lineages of Proteobacterial phytopathogens. The data rejected the null hypothesis and point to an evolutionary stasis of T3Es in mutualistic rhizobia.

RESULTS

Draft genome assemblies

We used paired-end Illumina sequencing to generate draft genome sequences for the selected strains of *S. fredii* and *B. japonicum* (Supplemental Table 2.1). Based on the following assessments, we determined that the assemblies and annotations were adequate for comprehensive mining of T3E genes (Pop and Salzberg, 2008; Klassen and Currie, 2012). The estimated genome sizes and predicted numbers of open reading frames (ORFs) were similar within each species and to finished reference sequences (Supplemental Table 2.1). Importantly, with a stringent criterion of ≥90% identity, 85.1% and 96.1% of the translated ORF sequences annotated in the draft USDA207 and USDA6 genome sequence, respectively, were similarly annotated in their corresponding genome sequences that were finished subsequent to our efforts (data not shown; Kaneko *et al.*, 2011; Margaret *et al.*, 2011). Analysis of aligned conserved regions, such as the T3SS-

encoding locus, indicated that the assemblies were of high quality and the draft genome sequences from USDA207 and USDA6 were similar in genome order to those of corresponding finished genome sequences (Supplemental Figure 2.1; data not shown; Kaneko *et al.*, 2011; Margaret *et al.*, 2011). Finally, the topology of a tree based on 624 translated orthologous sequences was similar to those generated using 16S or 23S rRNA-encoding regions and, in general, supported previous inferences of strain diversity (Supplemental Figure 2.2; van Berkum *et al.*, 2003). In total, these data indicate that the draft genome sequences were assembled and annotated to sufficient quality for a comprehensive investigation into the content and evolution of T3E genes in mutualistic bacteria.

Genome mining for candidate type III effector genes

Candidate T3E genes were identified based on their association with a *tts*-box, a *cis* element proposed to be recognized by TtsI, the two-component regulator-like protein that regulates T3SSs in rhizobia (Krause *et al.*, 2002; Marie *et al.*, 2004). We used a hidden markov model (HMM) to search the genomes and identified a total of 305 putative *tts*-boxes from the eight finished and draft genome sequences (Table 2.1). In *S. fredii*, 13~24 putative *tts*-boxes were identified from each of the three genome sequences, whereas in *B. japonicum*, upwards of 52 *tts*-boxes were found. In NGR234, other than the 11 *tts*-boxes previously found, we identified two additional *tts*-box sequences that were 1,345 bp and 1,119 base pairs (bp) upstream from NGR_a02270 (y4oB) and NGR_a00810 (*nodD2*), respectively (Marie *et al.*, 2004). As previously reported, no other putative *tts*-boxes were found in the chromosome of NGR234 or pNGR234b (Schmeisser *et al.*, 2009). A second T3SS-encoding locus is present on pNGR234b but its necessity in host infection has not been confirmed (Schmeisser *et al.*, 2009). In the finished genome sequence of USDA110, we found 52 *tts*-boxes, of which 29 were previously identified (Table 2.1; Zehner *et al.*, 2008). Fourteen of these *tts*-boxes are located upstream of 13 genes (bll1862 has two upstream *tts*-boxes) that encode proteins that are secreted in a T3SS-dependent manner (Zehner *et al.*, 2008).

However, our search failed to identify the other 10 previously identified *tts*-boxes, but none of the products encoded by the downstream genes appeared to be secreted in a T3SS-dependent manner (Zehner *et al.*, 2008).

To identify candidate T3E genes, we searched up to 10 kb downstream of the 305 *tts*-boxes. The reason for this relaxed criteria stems from observations that TtsI-regulated operons, such as the *nopB-rhcU* operon of NGR234, can be substantial in length (Perret *et al.*, 2003). We identified a total of 403 ORFs but were able to cull the list down to 277 (Table 2.1). The candidate T3E genes clustered into 96 different families based on sequence homology. As expected of T3Es and as a consequence of the filtering, 44% of the families had no members with matches to features in the NCBI conserved domain database (Marchler-Bauer *et al.*, 2011).

Functional testing of type III effectors for T3SS-dependent translocation

Proteins can be secreted or directly injected into host cells by the T3SS, the latter trait being the one distinguishing characteristic of a T3E. Due to the large number of candidates to be tested, we adopted the *Δ*79AvrRpt2 reporter in the γ-Proteobacterium *P. syringae* pv. tomato DC3000 (*Pto*DC3000) for high-throughput testing of candidate rhizobial T3E (Mudgett and Staskawicz, 1999; Guttman and Greenberg, 2001; Chang *et al.*, 2005). We first selected NopB and NopJ from NGR234 (NGR_a00680 and NGR_a02610, respectively) for validation of heterologous T3SS-dependent translocation. NopB is secreted *in vitro* in a flavonoid- and T3SS-dependent manner from NGR234, and NopJ is a member of the YopJ/HopZ T3E family (Ausmees *et al.*, 2004; Lorio *et al.*, 2004). The ORFs of *nopB* and *nopJ* were cloned downstream of a constitutive promoter and as translational fusions to Δ79AvrRpt2. The gene fusions were mobilized into *Pto*DC3000 and its T3SS-deficient mutant, *ΔhrcC*. Each of the strains were infiltrated into leaves of Arabidopsis ecotype Col-0 and examined for an HR approximately 20 hours post inoculation (hpi).

*Pto*DC3000 carrying a positive control, a fusion between the full-length *avrRpm1 P. syringae* T3E gene and *Δ79avrRpt2*, elicited a robust HR 20 hpi (Figure 2.1A; Debener

et al., 1991). Although Col-0 plants can elicit ETI in response to both AvrRpm1 and AvrRpt2, the HR we observed is known to be a consequence of perception of the latter by RPS2 since AvrRpt2, when co-delivered, "interferes" with AvrRpm1 perception (Dangl *et al.*, 1992). In contrast, *Pto*DC3000 lacking fusions to *Δ79avrRpt2* failed to elicit an HR but showed tissue collapse approximately 28 hpi, indicative of *Pto*DC3000-caused disease symptoms (data not shown). Importantly, *Pto*DC3000 carrying either the *nopB*::*Δ79avrRpt2* or *nopJ*::*Δ79avrRpt2* fusions elicited HRs within the same time frame and to the same robustness as the positive control (Figure 2.1A). The *ΔhrcC* mutant, regardless of the gene it carried, failed to elicit any phenotype throughout the course of the study, thereby confirming the T3SS-dependent delivery of T3Es. This is the first demonstration, to our knowledge, of heterologous T3SS-dependent delivery of rhizobial T3Es and confirmation of NopB and NopJ as T3Es. These data also validated the use of *Δ*79AvrRpt2 and *Pto*DC3000 for the rapid characterization of candidate rhizobial T3Es for T3SS-dependent translocation.

A total of 162 genes representative of the diversity of the 96 families were tested. From these, 74 T3Es belonging to 55 families were confirmed for T3SS-dependent translocation (Table 2.1; Figure 2.2). Prior to this study, 22 candidate T3E families had been previously predicted based on flavonoid-induced expression and T3SS-dependent *in vitro* secretion and an additional four proteins had been confirmed as *bona fide* T3Es based on *in vivo* translocation of fusions with the *cyaA* reporter from rhizobia cells (Schechter *et al.*, 2010; Wenzel *et al.*, 2010). Of the 25 candidate and confirmed families, we identified 22, of which 20 were validated as T3Es (Supplemental Figure 2.3). Forty-two of the candidate families had no evidence for T3SS-dependent translocation and were not considered T3Es based on the criteria used herein.

We retained the previously assigned Nop nomenclature (Nodulation Outer Proteins) for 17 families that were confirmed in this study (Marie *et al.*, 2001). The newly identified T3E families were assigned the names NopY through NopBT according to rules previously developed for naming T3E of pathogenic bacteria (Lindeberg *et al.*, 2005). A relational table of the validated T3E genes is provided that associates Nop names,

aliases, nucleotide and translated amino acid sequences, as well as other supporting information indicative of their classification as a T3E. Other than those previously identified, none of the translated sequences of the T3E genes identified in this study have detectable homology to proteins of known function (data not shown).

Conservation in content of type III effector collections

The draft genome sequences were distributed across many contigs, and physical gaps had the potential to both uncouple ORFs from their *tts*-box as well as fragment the coding sequences of T3E genes. As a consequence, the reliance on the HMM search of draft genome sequences could have caused us to miss some T3E genes. We therefore used sequences of members of confirmed T3E families in a homology-based approach to re-survey all draft genome sequences. Twenty additional homologs were identified and confirmed using locus-specific primers and PCR to amplify across the physical gaps to validate *in silico*-predicted gap closures. We also sequenced the PCR products of all fragmented T3E genes and all were considered functional based on the absence of premature termination codons (data not shown).

Among the T3E families, 10 of 18 and 31 of 42 T3E were "core" to the *S. fredii* and *B. japonicum* species, respectively. Furthermore, of these core T3Es, eight families (*nopB, nopM, nopP*, *nolU*, *nolV*, *nopY, nopZ* and *nopAA*) are distributed across both species (Figure 2.2). Only nine families were represented in just one strain. We used a reciprocal best BLAST hit approach to determine and compare the genome-wide orthology within and between species to that of the T3E families. Within *S. fredii*, the percentage of orthologous genes ranged from only 54.6% to 66.3% (Figure 2.3). Within *B. japonicum*, the percentage of orthologous genes ranged from 65.4% to 83.6%, which was similar to recent reports comparing the finished USDA6^{T} and USDA110 genome sequences (Figure 2.3; Kaneko *et al.* 2011). Therefore, the conservation in T3E content within the two species of rhizobia was similar to the estimated genome-wide conservation.

Low genetic diversity of rhizobial type III effectors

We calculated and plotted non-synonymous versus synonymous substitution rates (Ka/Ks) for the 43 confirmed T3E families that have more than one family member (Figure 2.4A). Of the 201 pairwise comparisons that could be examined, 95% had Ka/Ks values below 1, consistent with purifying selection on the T3E genes (Figure 2.4A). Only 10 pairwise comparisons within five T3E families had Ka/Ks values above 1, suggestive of positive selection. We also calculated Ka/Ks values for the *nod/fix* genes, which given their functions in nodulation and nitrogen fixation, are predicted to be under purifying selection, which was indeed the case (Figure 2.4B). Finally, we calculated Ka/Ks values for all orthologous genes encoded in the rhizobial genomes to determine how T3E families compared. Clustering of all predicted amino acid sequences from the eight strains resulted in 35,521 clusters, representing 65,694 orthologous pairs. Since 22,494 clusters only had a single representative, pairwise comparisons were calculated for orthologous pairs from 13,027 clusters (Figure 2.4C). As shown, most orthologous pairs had evidence for purifying selection and only 168 pairwise comparisons had Ka/Ks values greater than 1.0. This was expected, given the reported and observed high levels of identity between orthologous gene pairs within *S. fredii* and *B. japonicum* (Supplemental Figure 2.4; Kaneko *et al.,* 2011). Data suggest that T3E-encoding genes of *S. fredii* and *B. japonicum* are similar to the majority of genes in their genomes with evidence for purifying selection.

We also attempted to calculate Ka/Ks values for T3E genes of *P. syringae* for comparative purposes. However, results were not informative because ~33% of the families had only one member and another ~37% of the families had a member that was considered a pseudogene (data not shown; Baltrus *et al.*, 2011). This was in stark contrast to *S. fredii* and *B. japonicum*, where more than 50% of the possible comparisons representing 34 families were excluded because the nucleotide sequences were identical; in T3Es of *P. syringae* this was the case for less than 5% of the comparisons. Thus, these data underestimate the evidence for purifying selection for rhizobia T3Es and

underestimate the evidence for diversifying selection for *P. syringae* T3Es (Figure 2.4; data not shown).

For this reason, we quantified T3E sequence conservation/diversity in a different manner (Figure 2.5). Within the *S. fredii* and *B. japonicum* species, ~80% of the families had ≥ 90% amino acid identity among all family members. Twenty-one of the T3E families had all members with ≥ 99% identity and 10 families had within-family diversity less than 90% amino acid identity but were all nevertheless confirmed as T3Es. In the case of the NopB T3E family, we observed high levels of amino identity within a species, but low identity between species. The translated sequences of *nopB* of USDA207 and USDA257 are 100% identical to each other and have 98% identity to its member from NGR234. Within *B. japonicum*, the *nopB* members had ≥ 99% identity in all pairwise comparisons but only 32% identity to its family members from *S. fredii*. Regardless, as shown, each of the tested *nopB*::*Δ79avrRpt2* gene fusions were sufficient for *Pto*DC3000 to trigger an HR at 20hpi (Figure 2.1B). In contrast, no response was observed in plants infected with the *ΔhrcC* mutant carrying members of the *nopB* family (data not shown). Finally, only three families had pseudogenes as determined based on the presence of premature termination codons relative to other members. We provided a similar analysis of T3Es from five *P. syringae* pathovars of group I, including *Pto*DC3000 (Figure 2.5). As shown in bacteria that exemplify the co-evolutionary arms race, their T3Es show patterns distinctly different from that of *S. fredii* and *B. japonicum,* with fewer core T3Es, higher number of single T3Es and pseudogenes, and a far greater range of within-family amino acid sequence diversity.

Horizontal gene transfer

Given the relative lack of evidence for diversification of T3E genes or their loss via pseudogenization in *S. fredii* and *B. japonicum*, we asked whether there was evidence for T3E gene gain via horizontal gene transfer (HGT). Because of the challenges in using disjointed draft genome sequences to study HGT, we focused primarily on the finished *B.*

japonicum USDA110 genome. Furthermore, since all T3E genes of NGR234 are located on a plasmid, we decided against characterizing HGT in *S. fredii*.

The majority of the T3E genes clustered in the 680 kb symbiosis island that was previously identified in USDA110 (Kaneko *et al*, 2002; Göttfert *et al.*, 2001; Figure 2.6). This region tends to be detected based on differences in GC content relative to the rest of the genome. Nevertheless, there is no evidence for it to be mobile, unlike the corresponding Integrative and Conjugative element (ICE) in *M. loti*, and is likely a very "ancient" island (Kaneko *et al.*, 2002; Kaneko *et al.*, 2011; Göttfert *et al.*, 2001; Sullivan and Ronson, 1998; Ramsay *et al.,* 2006). A similar clustering of T3E genes to this symbiosis island was also observed in the genome of USDA6^{T} (data not shown; Kaneko et al., 2011). The few remaining T3E genes outside the island did not appear to associate to regions with signatures of HGT except for two near the origin of replication. In contrast, genes with evidence for positive selection, based on Ka/Ks values > 1, had a greater tendency to localize to regions with evidence for HGT (Figure 2.6). Finally, most of the conserved T3E genes, regardless of their location relative to the symbiosis island, shared similar genomic context between strains, potentially indicating their acquisition/presence in a common ancestor (Figure 2.7). A 25-kb genomic region of NGR234 and HH103, encoding genes of six T3E families, is entirely syntenous with only one neighboring gene showing any signs of variability (Figure 2.7A). Similarly, a large genomic region shared between USDA110 and USDA6^{T} is syntenous (Figure 2.7B). It therefore appears that most, if not all, T3E genes are core to the *S. fredii* and *B. japonicum* genomes.

DISCUSSION

Rhizobia are nitrogen-fixing bacteria comprised of nearly 100 different species. Many of these form mutualistic relationships with legumes that are crucial to the nitrogen cycle. We cataloged collections of type III effector genes from eight different strains within the *S. fredii* and *B. japonicum* species selected based on phylogenetic diversity as well as demonstrable T3SS-dependent host changes. We combined the use

of next generation sequencing to generate draft genome sequences and computational and experimental methods to identify T3E candidates and validate their T3SS-dependent delivery into plant cells, respectively. This study provides the first insights, to our knowledge, into the content of and selective pressures that shape collections of T3Es of mutualistic bacteria.

Evolution of type III effectors

The T3E collections of *S. fredii* and *B. japonicum* are highly conserved in both content and sequence, with very little evidence of diversifying selection, loss via pseudogenization, or acquisition via horizontal gene transfer (Figures 2.2, 2.4-2.7). Among the translocated T3E families of rhizobia, the majority of confirmed T3Es were "core" to the *S. fredii* and *B. japonicum* species. More than one-half of the families had members with greater than 97.5% identity in amino acid sequence and one-third of all families had 100% identity in sequence for all of its members. The observed evolutionary stasis is inconsistent with the arms race model of coevolution and therefore rejects the null hypothesis that collections of rhizobial T3Es evolve in a manner similar to those of Proteobacterial phytopathogens (Baltrus *et al.,* 2011; Stavrinides *et al.*, 2008; McCann and Guttman, 2008). Novelty in mutualism can result in instability. For this reason, mutualists may be under more pressure to limit diversification (Sachs *et al.,* 2011c). In this context, hosts select for the most frequent genotype of microbe, and the more common genotype of microbe is more likely to find a suitable host.

It could be argued that the apparent conservation of T3E genes in rhizobia simply mirrors the high overall conservation of their genomes. However, relative to plant pathogens, the genomes of *S. fredii* and *B. japonicum* had similar, if not lower, amounts of orthology among the strains characterized in this study (Figure 2.3; Baltrus *et al.*, 2011; Kimbrel *et al.*, 2010; Silby *et al.*, 2009; Kimbrel *et al.*, 2011). Furthermore, the T3E collections of pathogens are extremely diverse in spite of low genome diversity (Figure 2.5). For example, in *P. syringae*, two pairs of closely-related tomato pathovars DC3000 and T1, and the distantly-related cucumber pathovars, *lachrymans* 106 and 107, share

~80% and ~70% genome-wide orthology, respectively, yet have less than 50% conservation in their collections of T3Es (Baltrus *et al.*, 2011; Almeida *et al.*, 2009). Even when T3E family members are present, their flanking regions are not co-linear, suggesting that few extant T3E genes were present prior to the divergence of *P. syringae* (Chang *et al.*, 2005).

Many of the T3E genes of rhizobia are potentially members of operons that include genes encoding structural components of the T3SS. Furthermore, the majority of the highly conserved T3E-encoding genes are located in the symbiosis island where the *nod* and *fix* genes are also located. The conservation of the T3E genes may be a consequence of their association with functions essential to mutualism. Here again, a similar situation exists in *P. syringae*, but with a decidedly different outcome. T3E genes in the Conserved Effector Locus that borders the T3SS-encoding locus of *P. syringae* have pronounced nucleotide and functional diversity (*hopM1*), evidence for pseudogenization (*avrE1*, *hopM1*, *hopAA*) and presence/absence polymorphisms (*hopN*; Alfano *et al*., 2000; Baltrus *et al.,* 2011; Cai *et al.,* 2011). In contrast, even the rhizobial T3E genes distal to the symbiosis islands appear as static as those within the island with little evidence for HGT or sequence divergence (Figure 2.6). Taken together, the conservation in T3E collection content, the low number of unique genes, and evidence for purifying selection are consistent with the "mutualistic environment", where co-evolution of plant and rhizobia favor maintaining symbiosis (Sachs *et al.*, 2011; Law and Lewis, 2008).

Role of T3Es in mutualistic interactions

The ability of plants to recognize rhizobia and initiate PTI has been demonstrated, with PTI capable of reducing the efficiency of nodulation if elicited early during infection (Lopez-Gomez *et al.*, 2012). One hypothesis therefore posits that rhizobia use T3Es to manipulate host cells and dampen host PTI as an early step in colonization. The T3Es NopL and NopT of NGR234, for example, affect the MAPK defense signaling cascade and cause cytotoxic effects in transgenic plants, respectively, consistent with the idea that T3Es of rhizobia can perturb host defense (Yang *et al.*, 2009; Zhang *et al.*, 2011).

Additionally, candidate T3Es of rhizobia have been previously identified based on membership in families of T3Es of pathogens, suggesting, at the least, a similar molecular function (Dai *et al.*, 2008; Kambara *et al.*, 2009). In the one other case in which a T3SS has been characterized in the context of mutualism, a T3SS mutant of *Sodalis glossinidius*, a secondary endosymbiont, was severely compromised in infecting its insect host (Dale *et al.*, 2001).

Nevertheless, the hypothesis that rhizobia use T3Es to manipulate host cells and dampen PTI is difficult to reconcile in light of the repeated observations that the T3SS is seemingly not essential for rhizobial infection of all plants (Marie *et al.,* 2003; Dai *et al.*, 2008; Skorpil *et al.*, 2005; Hubber *et al.*, 2004). An alternative hypothesis suggests that the T3SS has more of an accessory or host-specific role rather than a necessary function. These data, in addition to the observation that not all rhizobia species encode for T3SS, suggest that T3Es, in particular, evolve to optimize interactions with specific host lineages. Arguing against this, however, is the observation that *S. fredii* NGR234 and USDA257 have exceptionally large host-ranges (Pueppke and Broughton, 1999). But, the extent to which the T3SS is required across these diverse species of host plants has not been explored.

T3Es and nodulation restriction

Rfg1 and *Rj2* are nodulation restriction genes of soybean that affect symbioses with *S. fredii* USDA257 and *B. japonicum* USDA122, respectively, in a T3SS-dependent manner (Yang *et al.*, 2010). Molecular characterization showed that both genes encode for NB-LRR motifs characteristic of R proteins. Hence, it logically extends that nodulation restriction is the same as ETI and that evasion of ETI likely explains the observation that T3SS-deficient mutants of rhizobia gain incompatible plants as hosts. In this scenario, loss of a functional T3SS would not only abrogate delivery of T3Es necessary for suppressing PTI, but also Avrs that can be perceived by R proteins to elicit ETI. Our cataloging and comparisons of T3E collections within the *S. fredii* and *B. japonicum* strains provides a very short list of candidates for the molecular demonstration of ETI in

legume-rhizobia interactions (Figure 2.2). Given the ability of *S. fredii* USDA207 to nodulate *Rfg1*-expressing plants, the likely Avr candidates are those polymorphic between the two *S. fredii* strains, i.e., NopBI, NopBJ, and NopBT.

Evasion of nodulation restriction via loss of T3SS may be associated with a cost, that being a compromise in the ability to effectively suppress PTI and a lower efficiency in establishing symbiotic interactions. In this regard, the T3SS and its delivered T3Es are implied as being necessary for full infection. As such, the ability of T3SS-deficient mutants to gain hosts is still consistent with the hypothesis that rhizobia use T3Es to manipulate host cells and dampen host PTI as an early step of colonization. Further, given that conservation can be interpreted as an indicator of functional importance, it is noteworthy that homology-based surveys of *B. japonicum* show the T3SS-encoding loci to be highly conserved across the strains that were characterized (Mazurier *et al.*, 2006).

An *R* gene-mediated defense response that limits nodulation by restricting a presumably mutualistic symbiosis seems counter-intuitive. One possible explanation is that ETI against rhizobia is simply due to convergence of pathogen and mutualist T3Es on defense proteins that are guarded by R protein(s). This guard hypothesis predicts that R proteins elicit ETI upon perception of changes to their guarded host protein (Jones and Dangl, 2006). Therefore, ETI against rhizobia is potentially an unfortunate consequence of rhizobia having to target guarded host proteins and the selective pressure for the host to maintain resistance against pathogens, rather than against beneficial rhizobium. Supporting this is the observation that nodulation restriction loci, such as *Rfg1*, are often clustered with large families of NB-LRR-encoding genes and genetically linked to resistance specificities against agricultural pests (Graham *et al.*, 2002).

Heterologous T3E translocation

Given the number of candidates tested and the length of time required for rhizobial-based assays, we elected to use a high-throughput heterologous system for testing (Schechter *et al.*, 2010; Wenzel *et al.*, 2010). Our analyses contributed T3E gene sequences from 52 different families and is the first demonstration, to our knowledge,

that *P. syringae* pv. *tomato* DC3000, a γ-proteobacteria, can deliver T3Es of rhizobia (Figure 2.1). We cannot exclude the possibility of false positives and false negatives, but the genes we classified as T3Es produced clear, robust, and repeatable HRs when assayed in *Pto*DC3000. Most importantly, our confirmed T3Es included nearly all candidates previously identified by others, suggesting that the employed assay did not lead to many false negatives (Supplemental Figure 2.3). On the other hand, we were unable to confirm the translocation of NopA and NopT, previously shown to be secreted in a type III-dependent manner (Deakin *et al.*, 2006). NopA may in fact be a structural component of the T3SS rather than a T3E (Deakin *et al.*, 2006). NopT, in contrast, is a member of the YopT/AvrPphB family and likely a *bona fide* T3E. The cytotoxic effects of NopT in Arabidopsis could have caused misleading conclusions in the translocation assay (Dai *et al.*, 2008). We did not test NopC, NopH or NopD for translocation because they were not identified or did not pass our filters (Rodrigues *et al.*, 2007; Hempel *et al.*, 2009; Deakin *et al.*, 2006).

The draft genome sequences could have also contributed to missing additional T3E genes. In fact, some of the predicted T3Es could not be tested as some were distributed across multiple contigs and recalcitrant to PCR amplification, which we inferred to be a result of local assembly mistakes. For example, the candidate T3E bll8244 from USDA110 was found in all *B. japonicum* strains as well as USDA207 and USDA257, but the short repeating sequences in this gene made it difficult to assemble (Pop and Salzberg, 2008). Next generation sequencing technology has experienced tremendous advances since we initiated this work and use of newer technologies could contribute to finishing the genome sequences. Regardless, given the sufficient quality of the draft genome sequences we used, we suspect that few TtsI-regulated T3E genes were missed and that their discovery would not likely impact the overall conclusions of this study.

Conclusion

It is evident that the type III secretion system is an important mechanism for bacteria to establish symbioses with their hosts, regardless of the outcome of the interactions. We mined eight strains from two species of rhizobia and identified a large number of T3E sequences, in which a substantial number were core to *S. fredii* and *B. japonicum*. While the mechanisms for these T3Es remain unresolved, it is apparent that the outcome of the symbiosis may influence the evolution of collections of T3Es, with pathogens in a co-evolutionary arms race and mutualistic rhizobia in a co-evolutionary armistice punctuated by infrequent skirmishes with their hosts.

ACKNOWLEDGEMENTS

We would like to thank an outstanding group of undergraduate researchers for their assistance (alphabetical order): Andres Alvarez, Philip Hillebrand, Denise Hrouda, Stanley Lee, Ryan Lilley, Meesha Peña, Liz Stoenner, Jayme Stout, and David Swader-Hines. We would also like to thank Mark Dasenko and Chris Sullivan in the CGRB for high-throughput genome sequencing and data preparation, Drs. Dan Arp, Michael Freitag, and Tom Wolpert for critical reading of this manuscript and insightful comments, as well as Dr. Peter van Berkum of the USDA ARS National Rhizobium Germplasm Collection for providing us the rhizobial strains. Finally, we thank Dr. Jeffery Dangl for his guidance, wonderful mentorship, and generosity in providing the space and resources to initiate this project.

MATERIALS AND METHODS

Bacterial strains and plasmids

Bacterial strains used in this study were: *B. japonicum* strains USDA6, USDA110, USDA122, USDA123, and USDA124; *S. fredii* strains USDA207 and USDA257; *Rhizobium* sp. NGR234, *Pseudomonas syringae* pv. *tomato* DC3000 (*Pto*DC3000), its T3SS-deficient mutant (Δ*hrcC*), and *Escherichia coli* DH5α. Rhizobia strains and *P. syringae* were grown in modified arabinose gluconate media (MAG) or King's B (KB) media, respectively, at

28°. *E. coli* DH5α was grown in Luria-Bertani (LB) media at 37°C. Antibiotics were used at the following concentrations: 50 μg/ml rifampicin (*Pto*DC3000), 30 μg/ml kanamycin (all bacterial strains), 50 μg/ml chloramphenicol (*B. japonicum* strains), and 25 μg/ml gentamycin (*E. coli*). Plasmids used in this work were pDONR207 (Invitrogen, Carlsbad, CA), the Gateway destination vector pDD62-Δ79AvrRpt2 (Mudgett *et al.*, 2000), and the pRK2013 conjugation helper plasmid (Figurski and Helinski, 1979).

Genome sequencing

Genomic DNA was extracted from *B. japonicum* strains USDA6, USDA122, USDA123, and USDA124, and *S. fredii* strains NGR234, USDA207 and USDA257 using osmotic shock, followed by alkaline lysis and phenol-chloroform extraction. We prepared 5 μg of DNA from each strain according to the instructions provided by the manufacturer (Illumina, San Diego, CA). Libraries were sequenced using the paired-end cycle sequencing kit on the Illumina (Supplemental Table 2.1). Sequencing was done by the Center for Genome Research and Biocomputing Core Labs (CGRB; Oregon State University, Corvallis, OR). Since the finished genome sequence for NGR234 was published early in our study, we used it rather than the draft genome sequence (Freiberg *et al.*, 1996; Schmeisser *et al.*, 2009).

Genome assembly and annotation

We used Velvet 0.7.55 to *de novo* assemble the genomes (Zerbino and Birney, 2008). Multiple assembles, using different parameters, were produced for each genome, and the highest quality assembly was identified using methods described previously (Kimbrel *et al.*, 2010). The Mauve Aligner 2.3 (default settings) program was used to order the contigs greater than 1 kb in length into scaffolds based on a closely related reference genome sequence (Rissman *et al.*, 2009). The reference sequence used for *B. japonicum* strains was USDA110, while the symbiotic plasmid of *Rhizobium* sp. NGR234 was used as a reference for the *S. fredii* strains (Freiberg *et al.,* 1997; Kaneko *et al.*, 2002). Genome assemblies were annotated using Xbase, and open reading frame (ORF)

annotations were further refined using the NCBI conserved domain database (CDD; Marchler-Bauer *et al.*, 2011; Altschul *et al.*, 1997; Lowe, 1997; Kurtz *et al.,* 2004; Delcher *et al.,* 2007; Lagesen *et al.*, 2007; Marchler-Bauer *et al.,* 2009).

Bioinformatic analyses

For visualization of genomic regions, BLAST atlases were generated using the Gview Server (Petkau *et al.*, 2010). USDA207 and USDA257 were compared against pNGR234a (NC_000914), and USDA6, USDA122, USDA123 and USDA124 were compared against USDA110 (NC_004463). ORFs longer than 100bp from pNGR234a or USDA110 (Supplemental Figure 2.5, black vertical lines) were used as queries against each genome, and hits with greater than 80% identity and an e-value ≤ $1e^{-15}$ were mapped against the genome position of the query sequence. BLAST Atlases were obtained through the Gview Server using BLASTN with the following settings: e-value ≤ 1×10^{-15}, alignment length ≥ 100 bp, and percent identity ≥ 80% [105]. The Circos was plot was generated using the Circos Table Viewer (Krzywinski *et al.*, 2009).

Non-synonymous (Ka) and synonymous (Ks) substitution rates were determined by first clustering all of the translated sequences using CD-HIT with (-c .75 -n 5 -T 6 -s 0.7) settings (Li and Godzik, 2006). We used *ad hoc* Perl shell scripts to call ClustalW2 to align the translated sequences, PAL2NAL to construct codon alignments from the ClustalW2 output, and KaKs_Calculator 1.2 with the 'YN' approximate method to calculate Ka/Ks values for all pairwise comparisons (Yang and Nielsen, 2000; Suyama *et al.*, 2006; Zhang *et al.*, 2006; Larkin *et al.*, 2007).

The Sanger Institute's Alien Hunter (default settings) was used to analyze genome sequences for potential horizontal gene transfer (HGT) events. Alien Hunter uses Interpolated Variable Order Motifs (IVOMs) to identify compositional biases and localize boundaries of predicted HGT regions (Vernikos and Parkhill, 2006).

T3E candidate discovery

We used sequences of 30 confirmed functional *tts*-boxes from *B. japonicum*, *S. fredii* and *M. loti* MAFF303099 to train a Hidden Markov Model (Zehner *et al.*, 2008; Marie *et al.*, 2004; Eddy, 1998; Sánchez *et al.*, 2009). An *ad hoc* Perl shell script was used to identify *tts*-boxes. The score of significance was calibrated to 5.0 based on the identification of 11 functionally validated *tts*-boxes located on the pNGR234a megaplasmid (Marie *et al.*, 2004). We next identified ORFs on the same strand as the *tts*-box, either up to 10 kb downstream or until another ORF on the opposite strand was encountered. We used BLASTX (e-value ≤ 1 x 10^{-15}) to filter out ORFs with translated sequences homologous to components of the T3SS, proteins encoded by organisms that lack a T3SS, or proteins with general housekeeping functions. We used BLASTN and sequences of candidate T3E-encoding genes to identify homologs from each of the eight genome sequences (e-value cutoff ≤1x10^{-15}).

T3Es were grouped into families based on BLASTP scores ≤ 1x10^{-15} across ≥50% the length of the protein. When all members of a family had amino acid identity ≥90% as determined using Clustalw, a single representative family member was chosen for testing (Larkin *et al.*, 2007). In families of < 90% amino acid identity, members representative of the diversity were tested.

T3E candidate cloning

Oligonucleotide primers were designed for candidate T3E-encoding genes to include a partial B1 and B2 sequence to the top and bottom oligonucleotide primers, respectively (sequences available upon request; Gateway® system (Invitrogen, Carlsbad, CA; Chang *et al.*, 2005). Gene-specific primers and 1 ul of genomic DNA were used in two-step PCR as previously described (Kimbrel *et al.*, 2011). PCR products were cloned into pDONR207 using BP clonase according to the instructions of the manufacturer (Invitrogen, Carlsbad, CA). PCR products were cloned into the destination vector pDD62-Δ79AvrRpt2 using LR clonase according to the instructions of the manufacturer (Mudgett

et al., 2000). Plasmids were transformed into *E. coli* DH5α cells and mated into *Pto*DC3000 or Δ*hrcC* via triparental mating.

***In planta* assay**

Arabidopsis thaliana Col-0 plants were grown in a controlled growth chamber environment (15-hour day at 22°C followed by 9-hour night at 20°C). *Pto*DC3000 cells were grown overnight in KB with appropriate antibiotics. Cells were washed and re-suspended in 10mM $MgCl_2$ at a final concentration of OD_{600}=0.1. These inocula were infiltrated into the abaxial side of leaves of ~6-week-old plants using 1-ml needle-less syringes. The hypersensitive response (HR) was scored approximately 20-24 hpi based on comparisons to *Pto*DC3000 carrying an empty vector or a fusion between *avrRpm1* and Δ79*avrRpt2*, negative and positive controls, respectively. Disease symptoms were scored approximately 28 hpi. Experiments were replicated a minimum of three times.

Table 2.1. Statistics for genome mining for T3E-encoding genes.

Strain*	# identified *tts*-boxes[†]	# ORFs post filtering[‡]	# ORFS post homology search[§]	T3Es[∫]	Not tested[#]
S. fredii NGR234*	13	24	24	12	2
S. fredii USDA207*	21	19	23	15	3
S. fredii USDA257	24	21	27	14	3
B. japonicum USDA6*	46	41	58	32	9
B. japonicum USDA110	52	51	66	36	10
B. japonicum USDA122	50	40	61	30	10
B. japonicum USDA123	47	39	60	36	7
B. japonicum USDA124	52	42	62	33	7

* The genome sequences of the (*) indicated strains were finished during various stages of this study (Schmeisser *et al.*, 2009; Weidner *et al.*, 2012; Kaneko *et al.*, 2011).

[†]A trained Hidden Markov Model (HMM) was used to identify candidate *tts*-boxes; the number of sequences with bit-scores ≥ 5.0 are presented.

[‡]ORFs within 10 kb and encoded on the same strand as the predicted *tts*-box were identified, and we used BLASTX (e-value ≤ 1 x 10-15) to filter out ORFs with translated sequences homologous to components of the T3SS, proteins encoded by organisms that lack a T3SS, or proteins with general housekeeping functions.

[§] The number of candidates after using BLASTN searches to identify additional homologies (e-value ≤ 1 x 10-15).

[∫] T3E-encoding genes based on T3SS-dependent elicitation of HR by *Pto*DC3000 in *Arabidopsis thaliana* Col-0.

[#]The number of candidate T3E-encoding genes that were recalcitrant to cloning.

Supplemental Table 2.1. Statistics for draft genome assemblies.

Strain*	Size (Mb) †	# of Contigs ‡	N50 (kb) §	Predicted ORFs ∫
S. fredii NGR234*	6.9	-	-	6322
S. fredii USDA207*	6.5	291	27.8	6995
S. fredii USDA257	7.0	384	26.7	6723
B. japonicum USDA6*	8.7	788	19.6	7960
B. japonicum USDA110	9.1	-	-	8317
B. japonicum USDA122	8.9	186	31	8107
B. japonicum USDA123	9.1	815	21.6	8361
B. japonicum USDA124	8.9	1285	19.3	7943

*The genome sequences of the (*) indicated strains were finished during various stages of this study.

†Sizes were estimated based on Velvet calculated values and included contigs from chromosomes and plasmids.

‡Only contigs greater than 1 kb in length were included; "-" = the finished genome sequences were used in this study.

§The contigs were ranked according to length and the contigs necessary to represent 50% of the estimated total size of the genome were identified. The size in kilobase (kb) of the smallest contig of the subset necessary to represent 50% of the genome is presented.

∫Draft genome sequences were annotated using XBase and the predicted numbers of ORFs larger than 150 bp in length are presented.

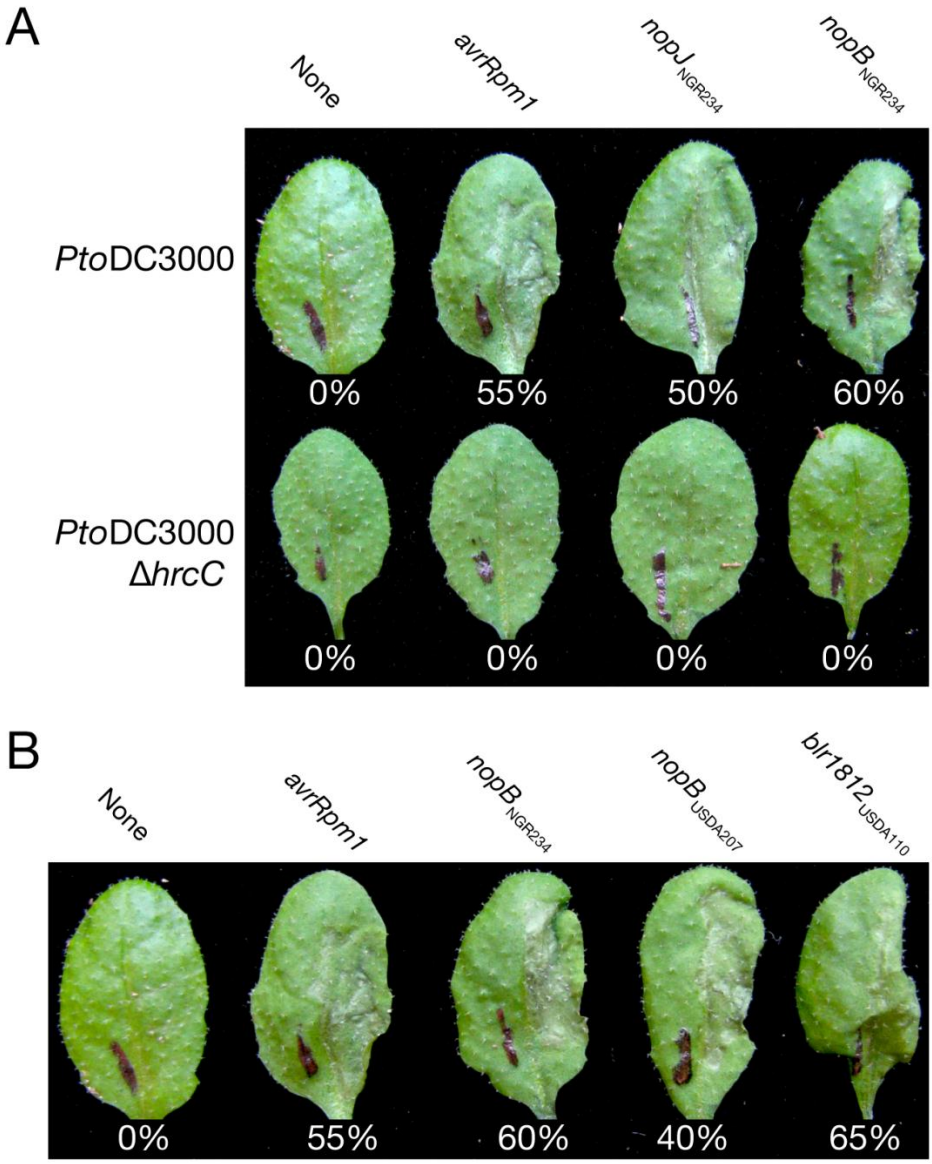


Figure 2.1. *Pto*DC3000 delivers T3Es of rhizobia in a T3SS-dependent manner.
(A) Leaves of Arabidopsis Col-0 (*Rps2*/*Rps2*) were infiltrated with *Pto*DC3000 (top row) and its T3SS-deficient mutant, *hrcC* (bottom row) carrying no gene fusion to Δ79*avrRpt2* or fusions to *P. syringae* T3E gene *avrRpm1* or NGR234 candidate T3E genes, *nopJ* or *nopB*. Members of the NopB T3E gene family all encode for functional T3Es. (B) Leaves of Arabidopsis Col-0 (*Rps2*/*Rps2*) were infiltrated with *Pto*DC3000 carrying no gene fusion to Δ79*avrRpt2* or fusions to *P. syringae* T3E gene *avrRpm1* or *nopB* genes from NGR234, USDA207, or USDA110. Leaves did not respond to infiltrations of *hrcC* (data not shown). In all experiments, leaves were scored for the HR ~20 hpi and the percent of responding leaves are presented (at least 20 leaves infiltrated). Experiments were repeated at least three times.

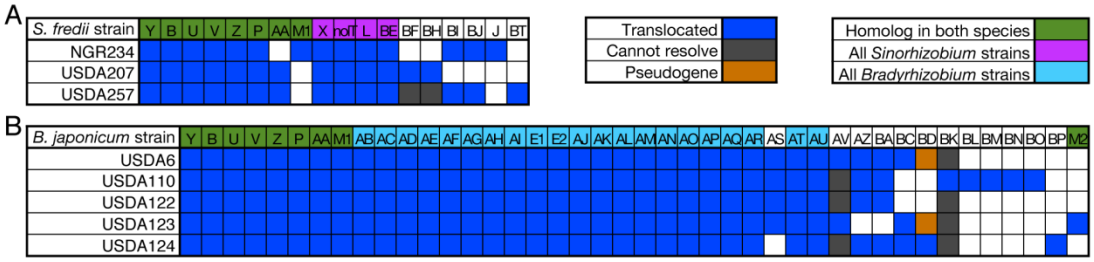


Figure 2.2. Distribution and conservation of T3E families in rhizobia.
The T3Es are listed across the top with strains of *S. fredii* (A) and *B. japonicum* (B) listed down the side. The conserved T3E-encoding genes are color-coded: between species (green) and within all three *S. fredii* (purple) or all five *B. japonicum* strains (cyan). The boxes are color-coded according to: functional T3E-encoding genes (blue); no evidence for T3SS-dependent delivery (red); homolog present but sequence could not be resolved (gray); homolog present with premature termination codon relative to other family members (brown); no detectable homolog (white).

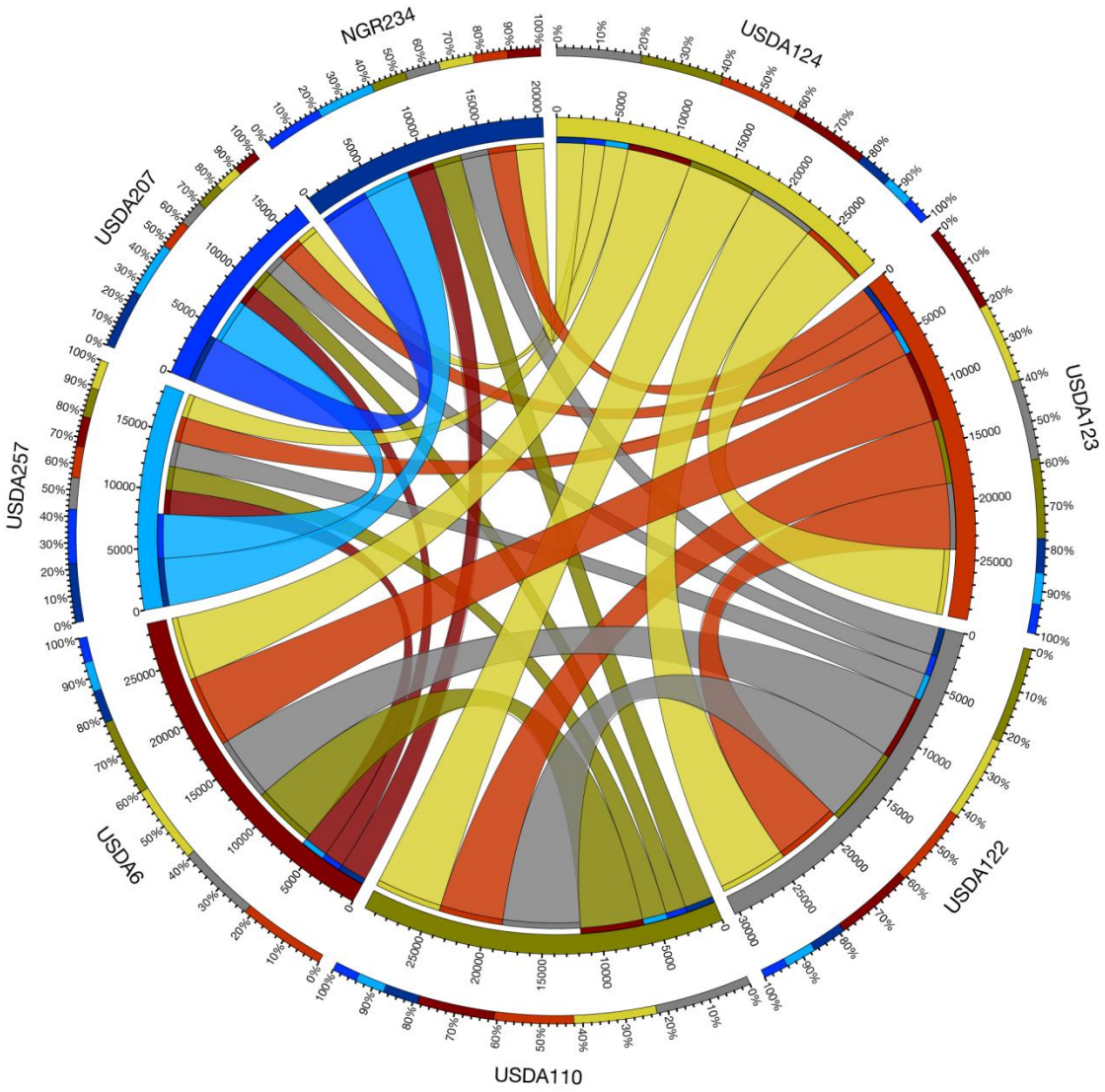


Figure 2.3. Circos visualization of genome-wide orthology within and between *S. fredii* and *B. japonicum* strains.

The most outer track, based on the length of the different colored bars, represents the amount of orthology other genomes have to the indicated genome. Genomes are ranked according to highest percent orthology (starting close to 0%) to lowest (ending closer to 100%). The inner track represents the amount of orthology the indicated genome has to the other genomes with interior ribbons connecting genomes, and variation in width depicting the extent of orthology. Genomes were assigned arbitrary colors (in a counter clockwise direction: NGR234 (dark blue); USDA207 (blue); USDA257 (cyan); USDA6 (maroon); USDA110 (olive); USDA122 (gray); USDA123 (orange); and USDA124 (yellow). Orthology was determined using reciprocal BLASTP of translated sequences between the eight rhizobia strains.

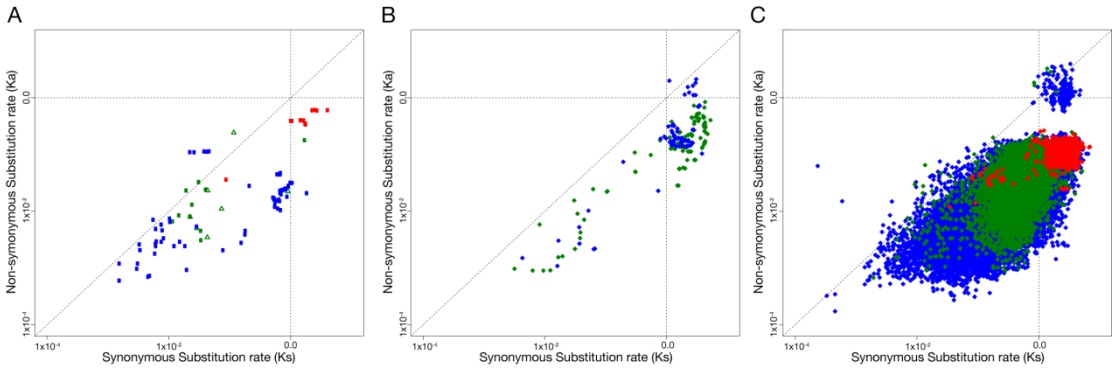


Figure 2.4. The majority of rhizobia T3Es have low Ka/Ks scores.
Synonymous (Ks) and non-synonymous (Ka) rates were plotted along the x- and y-axes, respectively, for all possible pairwise comparisons of rhizobia T3E-encoding genes (A), nod/fix genes (B), and all genes (C). Green, blue and red data points represent pairwise comparisons of genes within S*. fredii, B. japonicum*, or between species, respectively. For panel (A), squares and triangles represent family members that had both or only one member translocated, respectively. The dotted diagonal line indicates a ratio of 1. The dotted vertical and horizontal lines identify the boundaries for saturation of Ks and Ka, respectively. Comparisons between genes with identical nucleotide sequences or genes encoding potential pseudogenes were excluded from analyses.

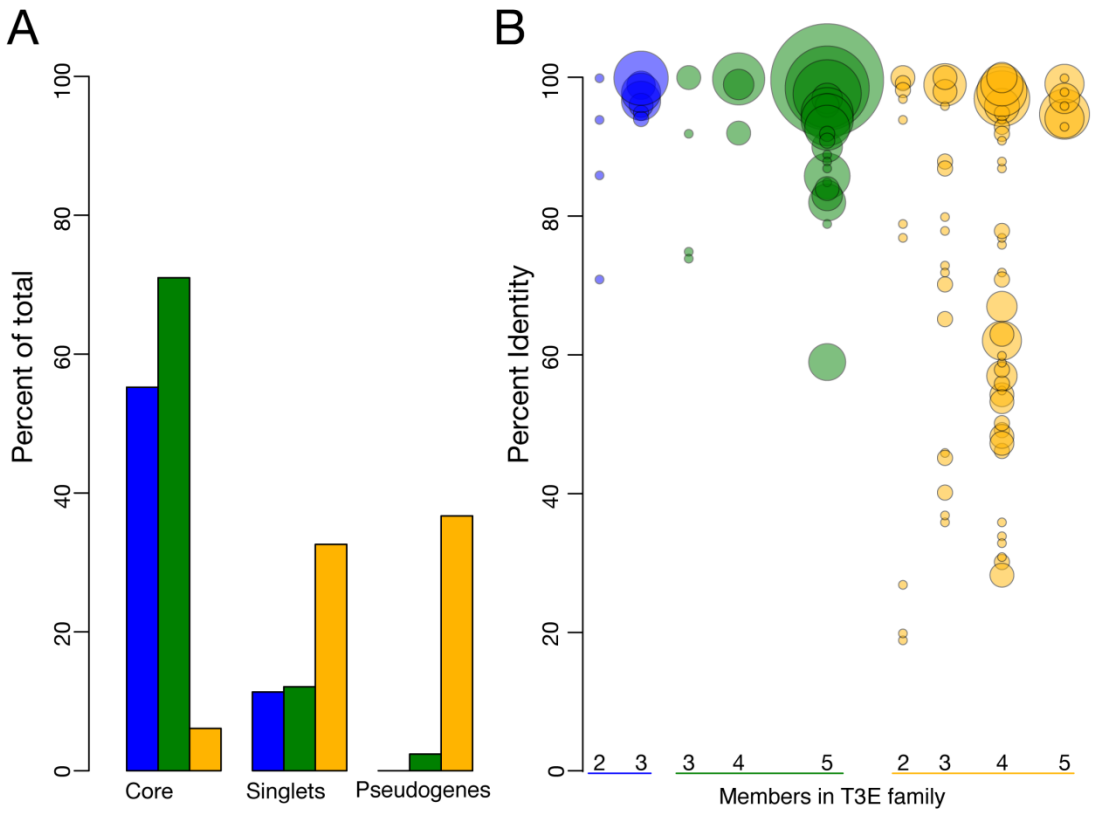


Figure 2.5. T3E gene conservation/diversity in *S. fredii*, *B. japonicum*, and Group I *P. syringae* pathovars.

(A) The percentage of core, singlets, and pseudogene T3Es were plotted based on the total number of families for *S. fredii* (18; blue), *B. japonicum*, (41; green), and group I *P. syringae* pathovars (49; gold). (B) Pairwise and within-family (according to number of members; x-axis) amino acid conservation was calculated for each T3E and plotted according to percent identity (y-axis). The number of pairs of T3Es with identical percent amino acid identity is represented by the size of the circle with the percent identity demarked by the center of each circle. Circles were drawn to scale except for those representing a single pairwise comparison.

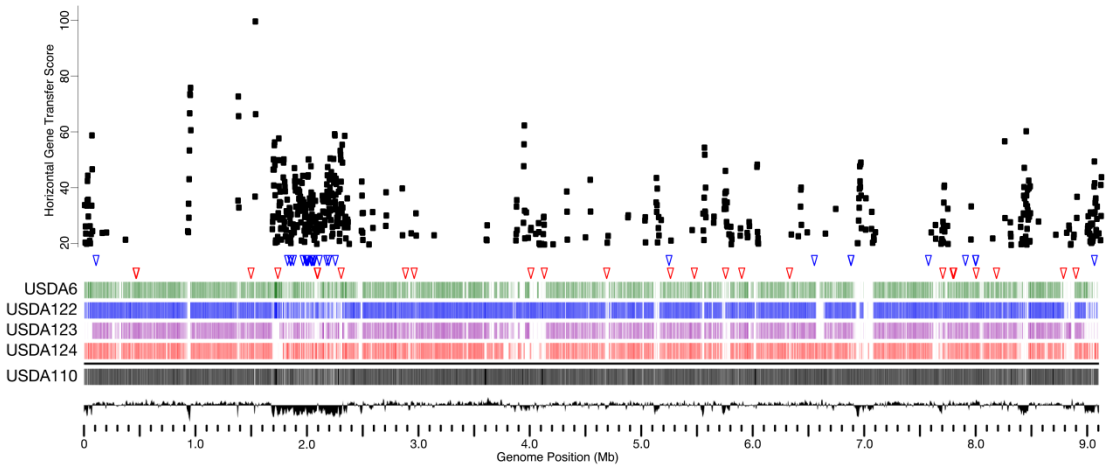


Figure 2.6. Analysis of *B. japonicum* genomes for evidence of HGT events.
(Top) The HGT scores as determined by Alien Hunter, are shown along the y-axis, only scores >20 are shown. The location of genes confirmed to encode T3E and those with Ka/Ks ratios > 1 of USDA110 are mapped according to the genome coordinates of USDA110 (blue and red triangles, respectively). (Bottom) BLAST Atlas showing orthology of *B. japonicum* genes > 100 bp in length (vertical lines) and plotted according to the genome coordinates of USDA110. Orthology was determined using BLASTN (e-value ≤ $1x10^{15}$; > 80% nucleotide identity). Bottom histogram depicts average GC% along a sliding window of 10 kb. The genome coordinates for USDA110 are presented along the x-axis in 100 kb increments; major ticks indicate megabase (Mb) increments.

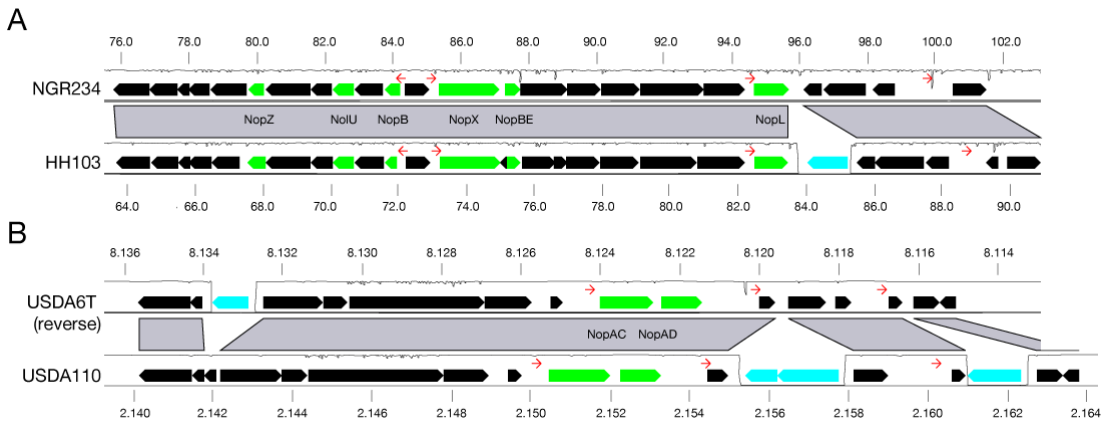
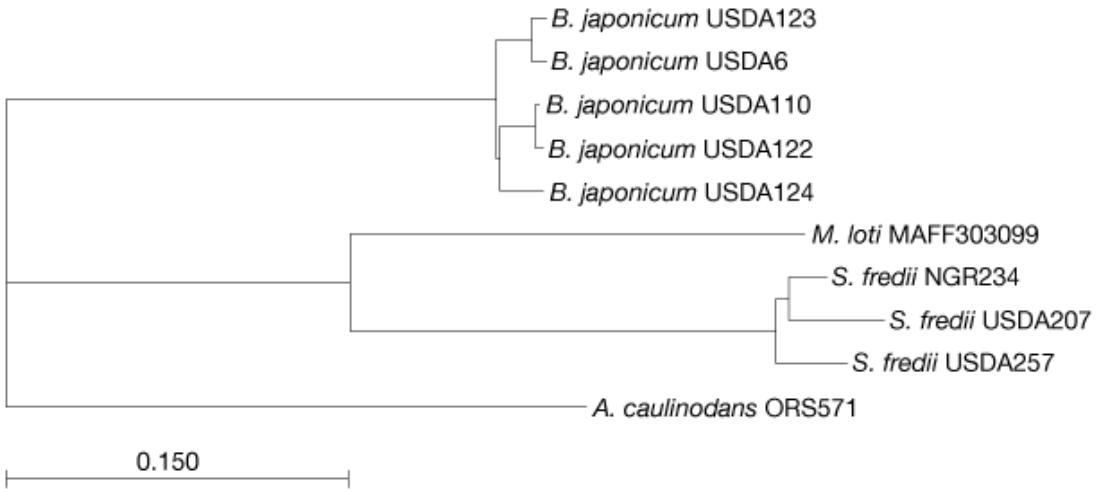


Figure 2.7. Synteny of representative conserved T3E-encoding genes.
A 25 kb region syntenous between pNGR234a and pSHH103d (A) and between USDA110 and USDA6^{T} (B). Syntenous blocks are highlighted as gray blocks with traces representing the amount of synteny as determined using Mauve. Thick arrows represent ORFs with direction indicating expression from the leading or lagging strand. Labeled and thick colored arrows depict candidate ORFs identified based on location relative to predicted *tts*-box (thin red arrows). Thick green arrows encode confirmed T3E. Thick black arrows represent syntenous ORFs, and thick cyan arrows represent non-syntenous ORFs. Numbers at the top and bottom indicate genome coordinates in kb or Mb for *S. fredii* and *B. japonicum*, respectively.

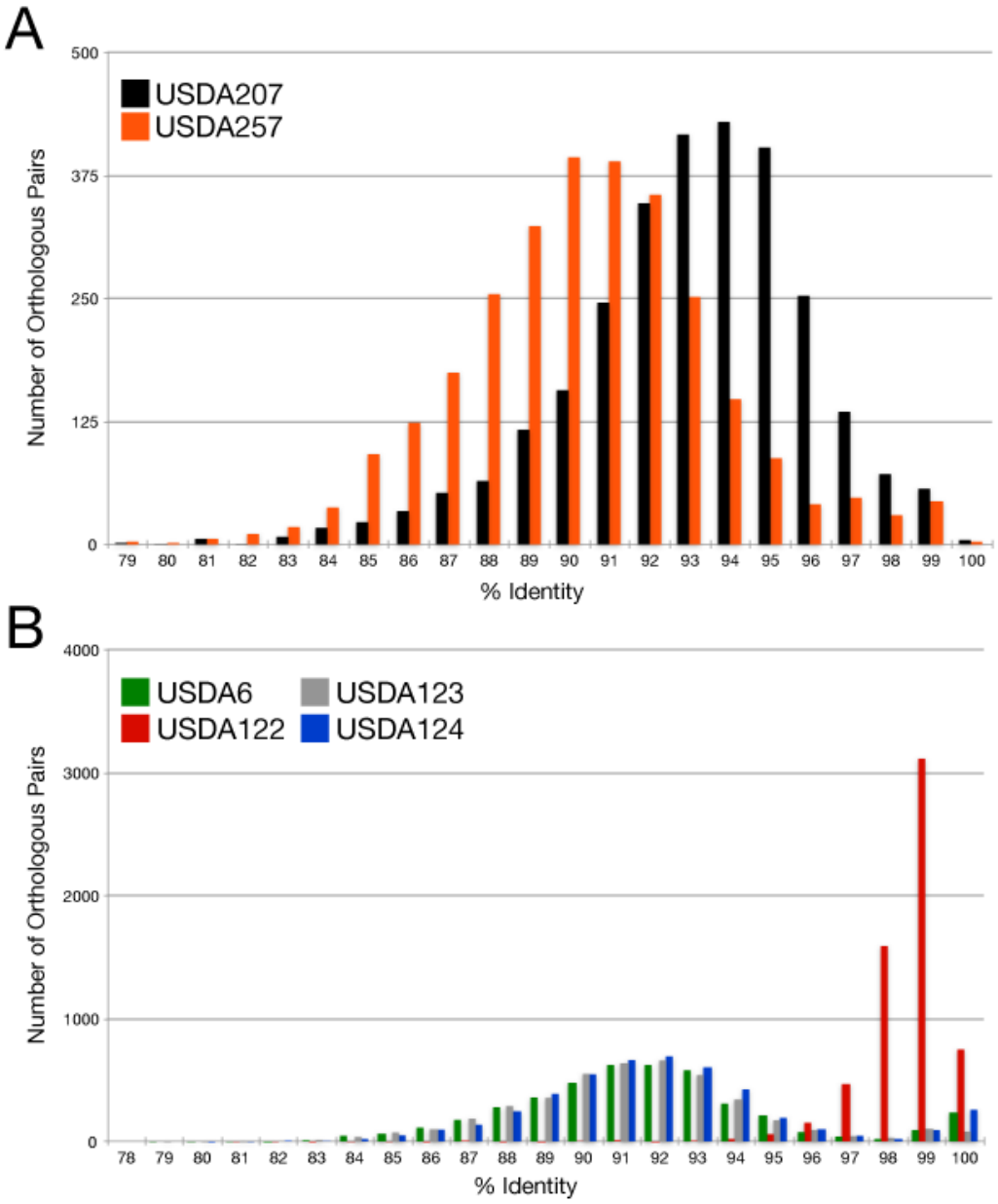
Supplemental Figure 2.1. Screenshot of Mauve alignment of T3SS-encoding loci of the eight strains of rhizobia.

Illumina short reads were *de novo* assembled and contigs were ordered using Mauve Aligner and NGR234 or USDA110 as a reference sequence (Freiberg *et al,.* 1997; Kaneko *et al.,* 2002). The T3SS-encoding loci were aligned using the outmost located T3SS-associated gene, NGR_a00520 (y4yS) as the landmark (thick blue vertical line). Large blocks represent syntenous regions are further indicated by the thin long connecting vertical lines, with the extent of synteny within each block depicted by the height of the colored regions. Mauve Aligner arbitrarily assigns colors. Short red vertical lines show contig breaks. Open boxes corresponding to NGR234 and USDA110 depict coding sequences. Coordinates are for each of the regions depicted (bp).

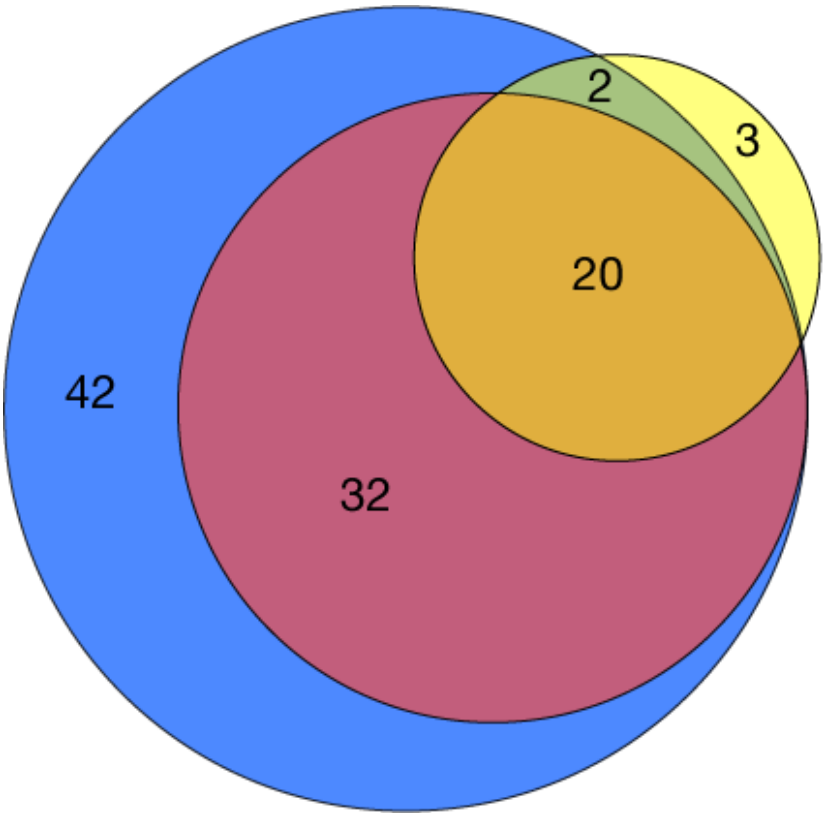


Supplemental Figure 2.2. Neighbor-joining phylogenomic tree of representative strains of rhizobia.

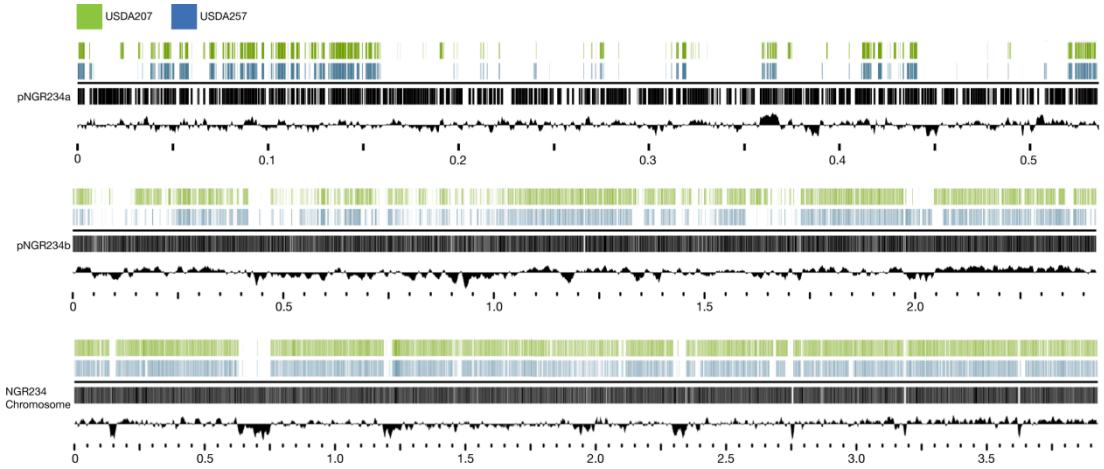
The tree is based on a super alignment of 624 translated amino acid sequences. *Azorhizobium caulinodans* ORS571 was included as the out group. Bootstrap values for each node are 100 (not shown); scale bar represents the number of substitutions per site. HAL (default settings) was used to generate whole-genome phylogenies of the six draft genome sequences and representative finished genome sequences (Robbertse *et al.,* 2011); NGR234 (NC_000914, NC_012586 and NC_012587; Freiberg *et al.,* 1997; Schmeisser *et al.,* 2009), USDA110 (NC_004463; Kaneko *et al.,* 2002), *Mesorhizobium loti* (NC_002678; Kaneko *et al.*, 2000), and *A. caulinodans* ORS571 (NC_009937.1; Lee *et al.,* 2008).



Supplemental Figure 2.3. Histogram of orthologs based on percent nucleotide identity. Orthologs between *S. fredii* (A) and *B. japonicum* (B) were identified using BLASTN (e-value ≤ $1x10^{30}$; > 90% the length of the coding sequence). For *S. fredii* ORFs from USDA207 and USDA257 were compared to the annotated ORFs in pNGR234a, pNGR234b and NGR234 chromosome (Freiberg et al., 1997; Schmeisser et al., 2009). For *B. japonicum*, the ORFs from the draft genomes of USDA6, USDA122, USDA123 and USDA124 were compared to the annotated ORFs of USDA110 (Kaneko et al., 2002). The nucleotide identity was determined for each orthologous pair, binned based on percent identity, and the numbers of each pair (y-axis) were plotted within each bin (x-axis).



Supplemental Figure 2.4. Area-proportional Venn diagram of candidate and confirmed rhizobial T3Es.
The yellow circle represents the union of 22 families of secreted and three families of delivered T3Es, previously discovered from rhizobia. Red and blue circles represent the 96 and 52 predicted and confirmed T3E families, respectively, identified in this study.



Supplemental Figure 2.5. BLAST Atlas of *S. fredii* genes.
Orthologs of > 100 bp in length (vertical lines) were determined using BLASTN (e-value ≤ $1x10^{15}$; > 80% nucleotide identity) and plotted according to the pNGR234a, pNGR234b, and chromosome coordinates (along the bottom). USDA207, USDA257, and NGR234 ORFs are represented by green, blue, and black, respectively, vertical lines. Histograms depict average GC% along a sliding window of 10 kb. Coordinates in Mb increments are presented.

Type III Effector Innovation in *Mesorhizobium loti* MAFF303099

William J. Thomas*, Jeffrey A. Kimbrel*, Ryan R. Lilley, and Jeff H. Chang

*These authors contributed equally to the work

Manuscript in preparation for *Molecular Plant-Microbe Interactions*

ABSTRACT

Rhizobia are a diverse collection of bacteria that can form mutualistic interactions with specific host plants. One determinant that affects host range of some rhizobia is the type III secretion system, a specialized delivery apparatus used by many Gram-negative, host-associated bacteria. This apparatus functions to deliver collections of bacterial-encoded type III effector proteins directly into host cells to modify host cell machinery and establish an environment conducive for infection. Type III secretion system loci have been identified from several species of rhizobia and their importance for establishment of infection is demonstrable. Nevertheless, the identity and function of type III effectors in mutualism remain uncharacterized. The lotus mutualist, *Mesorhizobium loti* MAFF303099 is unique among its species as being the only strain to employ a type III secretion system. As such, the characterization of MAFF303099 provides a rare opportunity to study the acquisition of type III effector genes. To this end, we employed a genome-enabled approach to mine the finished genome sequence of MAFF303099 for candidate type III effector genes. We employed a heterologous type III effector delivery system to catalog type III effectors based on the criterion of type III secretion system-dependent delivery. A total of 16 type III effector genes were identified and all were located in regions indicative of genome innovation, characterized by association with insertion sequences. The observations suggest that type III effector genes were not core to *Mesorhizobium* and were all recently acquired via horizontal gene transfer or innovated from existing genes.

INTRODUCTION

The rhizobia are a diverse collection of α-proteobacteria, many of which are able to form nitrogen-fixing nodules in compatible host genotypes. This fixation of nitrogen is a critical component of the nitrogen cycle, establishing a crucial role for the rhizobia-legume relationship in both natural ecosystems and agricultural environments. In order to successfully establish the interaction, the two symbiotic partners engage in a molecular dialog to establish the symbiotic relationship. Legume-specific flavonoids

induce the expression of nodulation (*nod*) genes in compatible rhizobia (Peck *et al.*, 2006; Jones *et al.*, 2007). The products of the *nod* genes synthesize Nod factors that are sufficient to cause morphological changes to compatible hosts, such as the formation of an infection thread and cortical cell activation (Spaink, 2000; Jones, 2010, Haney, 2011). The relationship culminates in the formation of nodules, specialized organs where differentiated bacteroids can fix atmospheric nitrogen.

Some species and strains of rhizobia, including *Bradyrhizobium japonicum*, *Sinorhizobium fredii*, *Rhizobium* NGR234, and *Mesorhizobium* loti MAFF303099, encode a type III secretion system (T3SS; Jiang and Krishnan, 2000; Göttfert *et al.*, 2001; Kaneko *et al.*, 2000). The T3SS is a specialized apparatus that functions to deliver collections of bacterial-encoded type III effector proteins (T3Es) directly into host cells. The T3SS genes are co-regulated with the *nod* genes, pointing to an early role in the establishment of symbiosis (Viprey, 1998; Hirsch, 2001; Krause, 2002; Wassem, 2008). Flavonoid perception elicits the *nodD* regulatory cascade, which regulates the expression of *ttsI*, a two-component regulatory-like encoding protein (Krause et al., 2002; Marie et al., 2004). TtsI is proposed to bind an upstream *tts*-box *cis* regulatory element that is typically upstream of genes encoding for components of the T3SS, candidate T3E, and other T3SS-associated accessory proteins (Zehner, 2008).

T3SSs and T3Es have been intensively studied in the context of plant-pathogen interactions. In this framework, T3Es collectively target and dampen MAMP-triggered immunity (MTI). This layer of defense is elicited upon perception of conserved microbe-associated molecular patterns (MAMPs) via plant pattern recognition receptors (PRRs), leading to a battery of induced immune responses (Segonzac and Zipfel, 2011). In the context of mutualism, the potential for MTI to limit host-rhizobial interactions has also been demonstrated and further supported by the inability of T3SS-deficient mutants to efficiently nodulate their once compatible host plants (Lopez-Gomez *et al*., 2012; Skorpil, 2005, Schechter, 2010; Wenzel, 2010; Yang, 2009). As demonstrated with pathogen T3Es, some rhizobial T3Es also have the potential to function in dampening defense, as NopL and NopT of *Rhizobium* NGR234 affect the MAPK defense signaling cascade and cause

cytotoxic effects in transgenic plants, respectively (Zhang, 2011). In total, these data suggest host MTI and the deployment of T3Es as a countermeasure against this barrier are components of the molecular dialog between host and mutualistic rhizobia.

T3Es, however, have the potential to betray the microbe to the plant through perception by resistance (R) proteins and their elicitation of another layer of defense called effector-triggered immunity (ETI). In the case of pathogens, perceived T3Es were classically called avirulence proteins because they rendered pathogens avirulent (Dangl and Jones, 2001). Striking parallels have also been observed in rhizobia, with substantial genetic and molecular data indicating that ETI also restricts nodulation by rhizobia in certain cultivars or species of plants. The ability of T3SS-deficient mutants of rhizobia to gain new hosts has been repeatedly observed and can be explained based on the failure of the T3SS mutants to deliver perceived avirulence proteins and evade R protein-mediated detection. In particular, *Rfg1* and *Rj2*, two nodulation restriction genes of soybean that affect symbioses with *S. fredii* USDA257 and *B. japonicum* USDA122, respectively, in a T3SS-dependent manner, encode for motifs classically associated with R proteins (Yang *et al.*, 2010).

T3E collections of plant pathogens have been modeled according to the co-evolutionary arms race. In this framework, plant and pathogen are under conflict and each partner evolves to maximize fitness at the expense of the other (Herre *et al.*, 1999; Douglas, 2008; Sachs *et al.*, 2010). The genetic patterns of pathogen T3Es provide strong support for this model (Ma and Guttman, 2008; Baltrus et al., 2011). In the plant pathogen *Pseudomonas syringae*, for example, examined strains exhibited extreme diversity in T3E content and sequence and are associated with genomic characteristics indicative of dynamic change, including hallmarks of horizontal gene transfer (HGT) and pseudogenization. The necessity to sufficiently dampen PTI, balanced by the need to avoid ETI, is thus a strong selective pressure that molds a microbe's collection of T3Es (Stavrinides, 2008, Cunnac, 2009; Cunnac, 2011). In stark contrast, T3E collections of *S. fredii* and *B. japonicum* have been modeled according to the competing mutualistic environment paradigm wherein both partners evolve to benefit the interaction (see

chapter 2; Kimbrel et al., in preparation). Indeed, the genetic patterns of rhizobial T3Es lend strong support for this model. The two species of rhizobia have large conserved core sets of T3Es with little sequence diversity and little evidence for gain via HGT or loss via deletion or pseudogenization (see chapter 2; Kimbrel et al., in preparation). These observations suggest that host and microbe are stabilized.

Mesorhizobium loti MAFF303099 requires a functional T3SS for efficient nodulation of its host, *Lotus* spp. (Sánchez *et al.*, 2009). Its T3SS locus is located within a large symbiosis island ubiquitous to *M. loti* strains that was potentially acquired as a mobile Integrative and Conjugative Element (ICE; Sullivan *et al.* 1995; Sullivan and Ronson, 1998; Kaneko, 2000; Sullivan *et al.*, 2002). Comparisons of the symbiosis island between MAFF303099 and *M. loti* R7A suggested it to be a hotspot for genome innovation. The backbone of the island represents only one-half of the sequence and is found as noncontiguous stretches interrupted by numerous insertions and deletions. Evidence suggests that the T3SS loci of MAFF303099 is not part of the backbone and in contrast, was a recent acquisition (Hubber *et al.*, 2004). Its integration was hypothesized to have resulted in the loss of the type IV secretion system locus (T4SS) from MAFF303099 since all other examined *M. loti* strains appear to encode for T4SS. Interestingly, one candidate T3E gene of MAFF303099, mlr6316, is a homolog of mis059 of R7A, which appears to be secreted by the T4SS (Hubber *et al.*, 2004; Sánchez et al., 2009).

MAFF303099 offers a unique opportunity to study the evolution of T3SS and, in particular, the evolution of T3Es. In this study, we used a genome-enabled approach to exhaustively catalog the T3E repertoire of MAFF303099. We mined the genome sequence of MAFF303099 to identify candidates and classified T3Es based on their location relative to a *tts*-box and direct delivery into host plant cells using a heterologous T3SS-dependent delivery, respectively (Kaneko *et al.*, 2000; Mudgett and Staskawicz, 1999; Guttman and Greenberg, 2001; Chang *et al.* 2005). Candidate T3Es had been previously identified based on their flavonoid-induced and T3SS-dependent secretion *in vitro*. In addition to the four previously identified candidate T3Es, we identified and

confirmed thirteen new T3E genes (Sánchez *et al*., 2009; Sánchez *et al*., 2012). All but two are located in the symbiosis island, and most are proximal to the T3SS locus. Two additional T3E genes were located outside of the symbiosis island on the pMLa plasmid. In all, these observations suggest that, unlike *S. fredii* and *B. japonicum*, no T3E genes are core to the *M. loti* genome and, like the T3SS loci, have also been recently acquired by MAFF303099 via HGT.

RESULTS AND DISCUSSION

Mining for candidate type III effector genes

We used a hidden markov model (HMM) to search the *M. loti* MAFF303099 genome for *tts*-boxes as a first step towards identifying candidate T3E genes. We identified 10 putative *tts*-boxes: nine were located on the chromosome, seven of which were found within the 611-kb symbiosis island (Table 3.1; Krause *et al.*, 2002; Marie *et al.*, 2004; Kaneko *et al.,* 2000). Six of the *tts*-boxes were found in a ~50-kb region containing the T3SS locus, with four upstream of T3SS-associated operons. The only *tts*-box distal to the T3SS-locus, designated Tb-a, was found on the 352-kb plasmid pMLa (Kaneko *et al.,* 2000). We also searched the corresponding symbiosis island of *M. loti* R7A for *tts*-boxes and found one sequence (Sullivan *et al.*, 2002).

The 10 *tts*-boxes we identified in MAFF303099 compare favorably to previous reports that described between seven and 27 *tts*-boxes for the MAFF303099 genome (Sánchez *et al*., 2009; Okazaki *et al.,* 2010). The seven *tts*-boxes described by Okazaki *et al.* (2010) corresponded to the top seven *tts*-boxes found by Sánchez *et al.* (2009), of which six met our bit-score threshold of 5.0. One of the *tts*-boxes detected by both previous studies did not meet our bit-score cutoff; nevertheless, we included it in our study based on evidence that it regulates a T3E (Sánchez *et al*., 2009; Okazaki *et al.,* 2010; Sánchez *et al.*, 2012). We did not identify the other 20 *tts*-boxes previously described, but whether they are *bona fide tts*-boxes remains unknown (Sánchez *et al.*, 2009). Indeed, three that were not detected by our screen also did not appear to respond to flavonoid induction, unlike reporter gene fusions to the *tts*-boxes associated

with *mlr6331, mlr6358, mlr6316,* and *mlr6361* (Sánchez *et al.*, 2009). The differences in the number of identified *tts*-boxes can be explained by the methods used as well as the sequences investigated. Sánchez *et al*. (2009) used only the eight invariant nucleotides distributed throughout the 30 nt-long HMM sequence to search for putative *tts*-boxes. Okazaki *et al.* (2010) did not report the methods used, but appear to have only focused on the 47.4 kb-long region encompassing the T3SS locus.

The number of *tts*-boxes present in *M. loti* MAFF303099 is low relative to other strains of rhizobia that we have screened. We identified nearly 25 *tts*-boxes each in *S. fredii* USDA207 and USDA257 and upwards of 50 *tts*-boxes in each of the five strains of *B. japonicum* (see Chapter 2; Kimbrel *et al*, in preparation). NGR234, by contrast, had only 13 such sequences, but all were found clustered on a small replicon, the 535 kb-long pNGR234a symbiosis plasmid (Schmeisser *et al.*, 2009).

To identify candidate T3E genes, we searched up to 10 kb downstream from each of the 10 *tts*-boxes to allow identification of candidate T3E genes potentially encoded in long operons (Perret *et al.*, 2003; see chapter 2; Kimbrel et al., in preparation). Thirty-five ORFs were identified; however, we culled 16 ORFs based on their translated sequences having homology to structural proteins of the T3SS or proteins with annotated functions common to non-T3E-encoding bacteria (Table 3.1).

Prior to our functional screen of candidate T3Es, there was already circumstantial evidence suggesting that at least half of the remaining 19 candidates were indeed T3Es. Seven of the MAFF303099 genes are orthologous to genes from other species of rhizobia confirmed to encode T3Es (see chapter 2; Kimbrel et al., in preparation). Four candidates unique to MAFF303099—Mlr6316, Mlr6331, Mlr6358, and Mlr6361—are secreted *in vitro*, in a flavonoid-inducible, T3SS-dependent manner (Sánchez *et al.*, 2009; Okazaki *et al.*, 2010; Sánchez *et al.*, 2012). Finally, MAFF303099 mutants of *mlr6316, mlr6358,* and *mlr6361* show a reduction in symbiotic fitness on compatible hosts and *mlr6361* has been demonstrated to have a role in restriction of nodulation on incompatible *Lotus* (Sánchez *et al.*, 2009; Okazaki *et al.*, 2010; Sánchez *et al.*, 2012).

Functional testing of T3Es for T3SS-dependent translocation

The translocation of T3Es directly into host cells via the T3SS is the one criterion that distinguishes T3Es from other T3SS-associated proteins. In order to test for direct delivery, we used the Δ79AvrRpt2 translocation marker and heterologous delivery from *Pseudomonas syringae* pv. *tomato* DC3000 (*Pto*DC3000) into cells of Arabidopsis (Mudgett and Staskawicz, 1999; Guttman and Greenberg, 2001; Chang *et al.,* 2005). This system was previously validated for rhizobial T3Es using NopB and NopJ from NGR234 (see chapter 2; Kimbrel *et al.*, in preparation). Briefly, when fusion proteins are delivered, the C-terminal portion of AvrRpt2 elicits ETI in an RPS2-dependent manner with a visible hypersensitive response approximately 20 hours post inoculation.

All of the 19 candidates from MAFF303099 were tested. Sixteen were delivered into plant cells, including the seven that belonged to families identified from *S. fredii*, and *B. japonicum* (Table 3.2; Kimbrel *et al.*, in preparation). In addition, NopBW, which we confirmed in this study, is homologous to a candidate T3E of *B. japonicum* that we could not previously confirm. The remaining nine T3Es of MAFF303099 are thus far unique to this strain. The newly classified rhizobial T3Es were named in accordance with the Nodulation Outer Protein (Nop) designations NopBG and NopBU to NopCB, as suggested by Marie *et al*., (2004) and based on established guidelines for T3E nomenclature in pathogenic bacteria (Lindeberg *et al.*, 2005).

Some of the MAFF303099 T3Es showed interesting homologies based on BLAST or PSI-BLAST searches. Both NopBV and NopBX (Mlr6331 and Mlr6361) are annotated as shikimate kinase-like proteins. Both also have homology to the *Ralstonia solanacearum* T3E Skwp2, and NopBV also has homology to the T3E XopAD found in two species of *Xanthomonas* (Bogdanove *et al.*, 2011). The rhizobial T3E NopBB (Mlr8765) contains a tetratricopeptide repeat (TPR) domain, which was described in the transcription activator-like (TAL) effectors of *Xanthomonas* (Scholze and Boch, 2011). PSI-BLAST searches identified an MlrC domain and a DUF1485 domain in NopBY, which is one of the two T3Es encoded within the symbiosis island but distal from the T3SS (Marchler-Bauer *et al.*, 2011). Interestingly, these two domains are also found in a protein of unknown

function, whose gene is located on the *S. fredii* NGR234 plasmid pNGR234b (Streit *et al.*, 2004). Finally, NopBZ shares a domain with the transcriptional regulator RpiR in *E. coli*, which regulates phosphosugar isomerases in that organism (Martinez-Cruz *et al.*, 2002; Marchler-Bauer *et al.*, 2011). It will be insightful to test whether these are functional domains and necessary for T3E function.

The genes we classified as T3Es produced clear, robust, and repeatable HRs when assayed in *Pto*DC3000. These confirmed T3Es included all four of the candidate T3Es predicted in previous work based on flavonoid-inducible, T3SS-dependent secretion (Sánchez *et al.*, 2009; Okazaki *et al.*, 2010; Sánchez *et al.*, 2012). Finally, in our cataloging of T3Es from eight species of *S. fredii* and *B. japonicum*, we convincingly showed that we identified and confirmed 20 out of 25 previously predicted candidate T3E families (see chapter 2; Kimbrel et al., in prepation). Nonetheless, given our reliance on a heterologous bacterium to deliver T3Es of rhizobia, we cannot exclude the possibility for false positives and false negatives.

T3E genes of MAFF303099 were likely acquired via horizontal gene transfer

Fourteen of the T3E genes clustered in the symbiosis island, while two were found on the 352-kb pMLa plasmid. In light of the evidence for the recent acquisition of the T3SS locus in MAFF303099, we asked whether there was evidence for horizontal gene transfer (HGT) events that led to the acquisition of T3E genes. We analyzed the entire chromosome and pMLa plasmid for signatures of HGT and correlated the signatures to the locations of the T3E genes (Figure 3.1). Furthermore, because of the high prevalence of insertion sequences in the MAFF303099 symbiosis island and their potential in shaping microbial genomes, we also correlated the location of these insertion sequences to the signatures of HGT and T3E genes (Figure 3.1).

Much of the 611-kb coding sequence of the symbiosis island was associated with many segmented areas with strong signatures for HGT. As such, it was difficult to associate evidence for HGT to any given T3E-encoding gene. This was not surprising given previous observations regarding the apparent dynamic nature of the symbiosis

island (Sullivan *et al.*, 2002). It is thus possible that the overall evidence for HGT masked any evidence for horizontal acquisition of T3E genes. Alternatively, the prevalence of HGT signatures may indicate an ongoing, persistent acquisition of new genes throughout most of the symbiosis island, in which the high incidence of insertion sequences may make it an evolutionary "cradle" for acquisition of new genes, including those encoding T3Es.

While four T3E genes were likely acquired by MAFF303099 *en bloc* with the T3SS locus as evidenced by their intimate association with the T3SS locus, eleven other genes may have been separately and independently acquired. Indeed, the mapping of insertion sequences provides support for such a model. A striking example of this is the *nopBY-nopBZ* operon, which is flanked by seven consecutive insertion sequences downstream and fifteen insertion sequences and five phage integrase/recombinases upstream of the operon. The acquisition of T3E genes through insertion sequence-mediated recombination events is common in pathogens; for instance, in the extreme case of the human pathogen *Shigella flexneri* 5a, over half of the ORFs present on a 210-kb virulence plasmid were related to insertion sequences, creating a mosaic of horizontally-acquired genes, including multiple T3E-encoding genes (Venkatesan *et al.,* 2001). The fish pathogen *Aeromonas salmonicida* subsp. s*almonicida* carries a small 18.5-kb plasmid with 20 ORFs, eight of which code for insertion sequences. Two of these insertion sequences flank a T3E gene and its cognate chaperone gene (Najimi *et al.*, 2009). Interestingly, two pathovars of the typically commensal bacterium *Pantoea agglomerans* appear to have recently acquired a plasmid-borne pathogenicity island carrying both a T3SS locus and a collection of T3E genes, all of which are associated with insertion sequences (Barash and Manulis-Sasson, 2007). Like MAFF303099, the *P. agglomerans* T3E genes associated with these insertion sequences have different origins: three of the six T3E genes have homologs in *P. syringae*, while the other three are unique to *P. agglomerans* (Barash and Manulis-Sasson, 2007).

Only two of the T3Es confirmed here are encoded outside of the symbiosis island. NopCA and NopCB are part of a putative short operon immediately downstream of the

tts-box designated Tb-a and found on the pMLa plasmid (Table 3.1). The *NopCA-NopCB* operon is found within a large cluster of hypothetical and unknown protein-coding genes. Interestingly, there is some evidence of recombination in this general region of the plasmid, as the operon is situated ~26 kb upstream of a cluster of integrase/recombinases, and ~28 kb downstream of an integrase/recombinase that appears to have disrupted a T4SS locus, leaving only *virD4* and *virB11.*

NopBG

NopBG (Mlr6316) is homologous to Msi059 of *M. loti* R7A, a protein that can be delivered by the type IV secretion system (Hubber *et al.*, 2004). Previous comparative analyses of the MAFF303099 and R7A symbiosis islands provided evidence that the loss of the T4SS in MAFF303099 either preceded or was caused by the acquisition of the T3SS (Sullivan *et al.*, 2002). This resulted in the deletion of the T4SS-encoding locus, as well as the *nod*-box upstream of the gene coding for the T4SS two-component regulator VirA. Interestingly, this recombination event inverted the sequence just downstream of the main T4SS-encoding locus, which included the *msi059* homolog *mlr6316* and a small piece of the T4SS-related gene *virD4*, such that *mlr6316* was relocated immediately upstream of the newly-acquired T3SS (Figure 3.2; Sullivan *et al.*, 2002). This relocation is particularly striking in the context of T3E innovation because it of the proximity of *mlr6316* to the *tts*-box designated Tb-1, located ~300 bp upstream of this gene (Table 3.1). It is unclear at this time whether the *tts*-box was already present or was later acquired through horizontal gene transfer, possibly mediated by the insertion sequence encoded just upstream of the *tts*-box; in either case, this represents a very complex instance of genetic innovation. Moreover, the T3E-encoding gene *mlr6316* of MAFF303099 and the T4E-encoding gene *msi059* of R7A show evidence of purifying selection based on our Ka/Ks analysis, despite their association with different secretion systems. Interestingly, our analysis indicates that a *tts*-box is present immediately upstream of the T4SS in the symbiosis island of R7A; indeed, it is the only *tts*-box we identified in the available R7A sequence. It will be interesting to test whether the *tts*-box

in R7A can be recognized by TtsI of MAFF303099 and whether msi059 can be delivered in a T3SS-dependent manner.

Concluding remarks

Using a genome-enabled computational screen for *tts*-boxes coupled with a functional screen for T3SS-dependent translocation, we identified 16 T3Es of MAFF303099, eight of which represent novel Nop families. All but two of the T3Es identified in this study were associated with the 611-kb symbiosis island encoded on the chromosome. This symbiosis island appears to be an evolutionary hot spot in which an unusually high concentration of diverse and active insertion sequences mediates HGT via integration of exogenous DNA (Uchiumi, 2004). HGT appears to be responsible for the acquisition of a T3SS locus by MAF303099 relatively recently in comparison to other T3SS-encoding rhizobia such as *S. fredii* and *B. japonicum*. Based on the work herein, it appears that four T3E genes were acquired along with the T3SS locus. It is likely that 11 other T3E genes were acquired independent of the T3SS loci, also likely a result of HGT given the high incidence of flanking IS sequences. Finally, one T3E gene, *nopBG* may have been present prior to the divergence of MAFF303099 from other *M. loti* species, and as suggested by Sullivan *et al.* (2002), potentially a T4E.

In pathogenic bacteria, collections of T3Es have been modeled as a co-evolutionary arms race with their hosts, as opposed to T3Es of *S. fredii* and *B. japonicum*, which we have modeled according to the mutualistic environment model of a co-evolutionary armistice. Although the genetic patterns of the T3Es of MAFF303099 contrast with the patterns we observed in other rhizobia, and are more similar to the genetic patterns of pathogen T3Es, we do not believe that MAFF303099 is in a co-evolutionary arms race. Rather, given the potential that MAFF303099 recently acquired its T3SS locus, we hypothesize that this strain is still relatively early in the molding of its collection of T3E genes and over time will evolve to have a stable collection of T3Es.

MATERIALS AND METHODS

Bacterial strains and plasmids

Bacterial strains used in this study were: *Mesorhizobium loti* MAFF303099, *Pseudomonas syringae* pv. *tomato* DC3000 (*Pto*DC3000), its T3SS-deficient mutant (Δ*hrcC*), and *Escherichia coli* DH5α. Rhizobia strains and *P. syringae* were grown in modified arabinose gluconate media (MAG) or King's B (KB) media, respectively, at 28°. *E. coli* DH5α was grown in Luria-Bertani (LB) media at 37°C. Antibiotics were used at the following concentrations: 50 μg/ml rifampicin (*Pto*DC3000), 30 μg/ml kanamycin (all bacterial strains), 50 μg/ml chloramphenicol (*B. japonicum* strains), and 25 μg/ml gentamycin (*E. coli*). Plasmids used in this work were pDONR207 (Invitrogen, Carlsbad, CA), the Gateway destination vector pDD62-Δ79AvrRpt2 (Mudgett *et al.*, 2000), and the pRK2013 conjugation helper plasmid (Figurski and Helinski, 1979).

Bioinformatic analyses

The Sanger Institute's Alien Hunter (default settings) was used to analyze genome sequences for potential horizontal gene transfer (HGT) events. Alien Hunter uses Interpolated Variable Order Motifs (IVOMs) to identify compositional biases and localize boundaries of predicted HGT regions (Vernikos and Parkhill, 2006).

T3E candidate discovery

We used sequences of 30 confirmed functional *tts*-boxes from *B. japonicum*, *S. fredii* and *M. loti* MAFF303099 to train a Hidden Markov Model (Zehner *et al.*, 2008; Marie *et al.*, 2004; Eddy, 1998; Sánchez *et al.*, 2009). An *ad hoc* Perl shell script was used to identify *tts*-boxes. The score of significance was calibrated to 5.0 based on the identification of 11 functionally validated *tts*-boxes located on the pNGR234a megaplasmid (Marie *et al.*, 2004). We next identified ORFs on the same strand as the *tts*-box, either up to 10 kb downstream or until another ORF on the opposite strand was encountered. We used BLASTX (e-value ≤ 1×10^{-15}) to filter out ORFs with translated

sequences homologous to components of the T3SS, proteins encoded by organisms that lack a T3SS, or proteins with general housekeeping functions.

T3E candidate cloning

Oligonucleotide primers were designed for candidate T3E-encoding genes to include a partial B1 and B2 sequence to the top and bottom oligonucleotide primers, respectively (sequences available upon request; Gateway® system (Invitrogen, Carlsbad, CA; Chang *et al.*, 2005). Gene-specific primers and 1 ul of genomic DNA were used in two-step PCR as previously described (Kimbrel *et al.*, 2011). PCR products were cloned into pDONR207 using BP clonase according to the instructions of the manufacturer (Invitrogen, Carlsbad, CA). PCR products were cloned into the destination vector pDD62-Δ79AvrRpt2 using LR clonase according to the instructions of the manufacturer (Mudgett *et al.*, 2000). Plasmids were transformed into *E. coli* DH5α cells and mated into *Pto*DC3000 or Δ*hrcC* via triparental mating.

***In planta* assay**

Arabidopsis thaliana Col-0 plants were grown in a controlled growth chamber environment (15-hour day at 22°C followed by 9-hour night at 20°C). *Pto*DC3000 cells were grown overnight in KB with appropriate antibiotics. Cells were washed and re-suspended in 10mM $MgCl_2$ at a final concentration of OD_{600}=0.1. These inocula were infiltrated into the abaxial side of leaves of ~6-week-old plants using 1-ml needle-less syringes. The hypersensitive response (HR) was scored approximately 20-24 hpi based on comparisons to *Pto*DC3000 carrying an empty vector or a fusion between *avrRpm1* and Δ79*avrRpt2*, negative and positive controls, respectively. Disease symptoms were scored approximately 28 hpi. Experiments were replicated a minimum of three times.

Acknowledgements

We would like to thank Philip Hillebrand and Liz Stoenner for their assistance and Dr. Peter van Berkum of the USDA ARS National Rhizobium Germplasm Collection for

providing us *M. loti* MAFF303099. We also thank Dr. Jeffery Dangl for his guidance, wonderful mentorship, and generosity in providing the space and resources to initiate this project. This work was supported by the National Research Initiative Competitive Grants Program Grant no. 2008-35600-04691 to JHC.

Table 3.1. *tts*-boxes identified in *M. loti* MAFF303099

***tts*-box**	**Sequence**[ʊ]	**Bit score***	**Coordinates**[ǂ]	**ORFs**[Ᵽ]
Tb-1	ccGTCAGcTtgtCGtaAGCcaacccgcctA	20.1	5125449…5125478	*mlr6316*, *mlr6318*
Tb-2	tcGTCAGtTtacCGaaAGCtaaaccgctcA	18.4	5153852…5153881	*mll6337*
Tb-3	caGTCAGcTtgtCGtcAGCtcggccacctA	21.8	5154733…5154762	*nopB*[§], *nolT*[§], *nolU*[§], *nolV*[§], *hrcN*[₡], *nopZ*[§], *hrcQ*[₡], *hrcR*[₡], *hrcS*[₡], *hrcT*[₡], *hrcU*[₡]
Tb-4	tcGTCAGgTtctCGaaAGCtcctgctcgtA	18.7	5164112…5164141	*nopC, nopA, nopY*[§], *hrcV*[₡], *mlr8765*
Tb-5	gcGTCAGcTcacCGtcAGCtcgttcgagtA	14.9	5172394…5172423	*mlr6358*
Tb-6	ccGTCAGcTaatCGtcAGCcaagcgatctA	12.6	5176483…5176512	*mlr6361*
Tb-7	tgGTCtGcTtgtCGacAGCctggtccatgA	7.2	619821…619850	*mll0786, mll0787, mll0788*
Tb-8	tcGCCAGcTggtCGacgGCaaagccgatcA	10.1	3937148…3937177	*mlr4966, mlr4967, mlr4968*
Tb-9	ccGTCAGcagatgGaccGgttggctttctA	6.0	4844156…4844185	*mlr5988*, *mlr5989*, *mlr5990*,
T-10	agGTCAGagtcaCGtcAGgtgaaggggatA	2.7[₹]	5139478…5139507	*mll6331*
Tb-a*	ccGcCAGcTtgtCGgcAGCcgccctcgatA	7.0	264871…264900	*mlr9294, mlr9295*, *mlr9297*, *mlr9300*

*Sequences with bit scores ≥5 were called *tts*-boxes

[ʊ]Invariant nucleotides are capitalized

[ǂ]Chromosomal location of *tts*-box (with the exception of Tb-a, located on pMLa plasmid

[Ᵽ]Underlined ORFs are candidate T3Es tested for translocation

[§]Confirmed T3E in other rhizobia species (see Chapter 2; Kimbrel *et al.*, in preparation)

[₡]T3SS structural protein

[₹]Bit-score did not meet criterion, but was considered a hit based on literature (Sánchez *et al.*, 2009; Okazaki *et al.*2010)

Table 3.2. Confirmed T3Es of *M. loti* MAFF303099.

Nop family	Gene ID	*tts*-box	Homologs
nopB	*mlr8763*	Tb-3	*S. fredii, B. japonicum*
nolU	*mlr8764*	Tb-3	*S. fredii, B. japonicum*
nolV	*mlr6341*	Tb-3	*S. fredii, B. japonicum*
nopX	*mll6337*	Tb-2	*S. fredii, B. japonicum*
nopY	*mlr6347*	Tb-4	*S. fredii, B. japonicum*
nopZ	*mlr6343*	Tb-3	*S. fredii, B. japonicum*
nopBB	*mlr8765*	Tb-4	*S. fredii**
nopBG	*mlr6316*	Tb-1	
nopBU	*msr6318*	Tb-1	
nopBV	*mlr6331*	Tb-10	*skwp2, skwp3, skwp4, xopAD*; shikimate kinase-like
nopBW	*mlr6358*	Tb-5	*B. japonicum*, hlk1, hlk2*
nopBX	*mlr6361*	Tb-6	*skwp2*, *skwp3, skwp4;* shikimate kinase-like
nopBY	*mlr5988*	Tb-9	DUF1485, MlrC C-terminal
nopBZ	*mlr5989*	Tb-9	RpiR-like protein
nopCA	*mlr9295*	Tb-a	
nopCB	*mlr9300*	Tb-a	DUF2130 superfamily

*Translocation has not been confirmed in this species

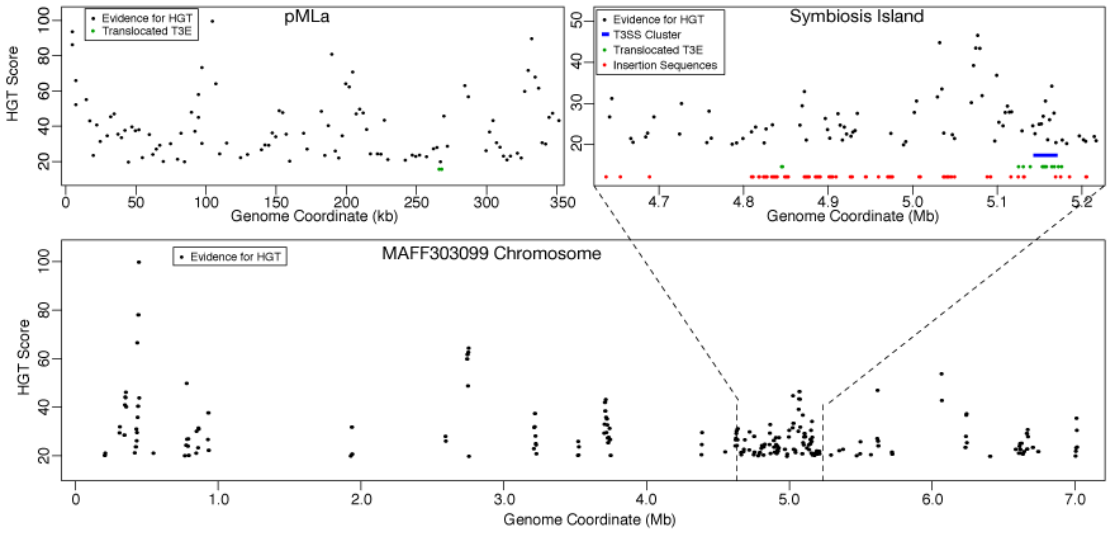


Figure 3.1. Analysis of *M. loti* MAFF303099 genome for evidence of HGT events.
HGT scores as determined by Alien Hunter are shown on the y-axis, plotted against the coordinates of the genome (bottom) and the pMLa plasmid (top left). The inset at top right represents the region encoding the symbiosis island. The location of confirmed T3E-encoding genes and insertion sequences are indicated on the x-axis with green and red dots, respectively, while the T3SS is designated by the blue bar.

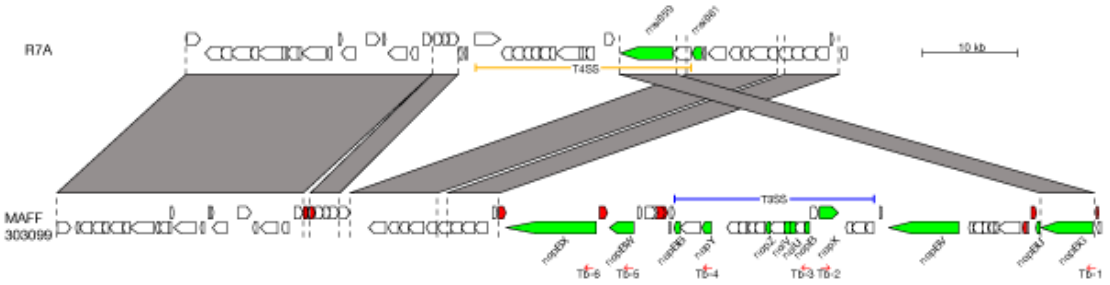


Figure 3. 2. Comparison of T4SS- and T3SS-encoding regions in *M. loti* MAFF303099 and R7A.

The regions including the T4SS (yellow line) and T3SS (blue line) loci of *M. loti* R7A and MAF303099 are shown along the top and bottom, respectively. Syntenous regions are indicated by the gray blocks and demarked by vertical hashes. Coding sequences are presented as open boxes with directions indicated. Green filled boxes represent candidate T4E or confirmed T3E genes and red filled boxes represent insertion sequences. The *tts*-boxes are indicated by red arrows. The sequence of the MAFF303099 region shown is reversed.

Recombineering and stable integration of the *Pseudomonas syringae* pv *syringae* 61 *hrp/hrc* cluster into the genome of the soil bacterium *Pseudomonas fluorescens* Pf0-1

William J. Thomas, Caitlin A. Thireault, Jeffrey A. Kimbrel, and Jeff H. Chang

The Plant Journal
2009 vol. 60 (5) pp. 919-928
PMID: 19682294

SUMMARY

Many Gram-negative bacteria use a type III secretion system (T3SS) to establish associations with their hosts. The T3SS is a conduit for direct injection of type III effector proteins into host cells where they manipulate the host for the benefit of the infecting bacterium. For plant-associated pathogens, the variations in number and amino acid sequences of type III effectors as well as their functional redundancy make studying type III effectors challenging. To mitigate this challenge, we developed a stable delivery system for individual or defined sets of type III effectors into plant cells. We used recombineering and Tn5-mediated transposition to clone and stably integrate, respectively, the complete *hrp/hrc* region from *Pseudomonas syringae* pv *syringae* 61 into the genome of the soil bacterium *Pseudomonas fluorescens* Pf0-1. We describe our development of Effector-to-Host Analyzer (EtHAn) and demonstrate its utility for studying effectors for their *in planta* functions.

INTRODUCTION

Pathogens must overcome plant defenses in order to successfully infect their hosts. One defense mechanism that plant pathogens encounter is basal defense, or PAMP-triggered Immunity (PTI). PTI relies on pattern recognition receptors (PRRs) to perceive microbes by their conserved molecular patterns (pathogen- or microbial-associated molecular patterns; PAMPs or MAMPs, respectively, Ausubel, 2005; Jones and Dangl, 2006; Schwessinger and Zipfel, 2008). Perception results in the induction of plant responses that include the deposition of callose into the cell walls (Schwessinger and Zipfel, 2008). From hereafter for the sake of simplicity, we use the terms PAMPs in reference to both PAMPs and MAMPs and PTI in reference to basal defense regardless of the source of the molecular pattern being perceived by the plant.

Many Gram-negative phytopathogenic bacteria deliver type III effectors into host cells through a type III secretion system (T3SS) as countermeasures against PTI (Cunnac *et al.*, 2009). The more than 20 genes encoding the regulatory elements and structural components of the T3SS are often clustered together in bacterial genomes (Galan and

Wolf-Watz, 2006). In *Pseudomonas syringae* the T3SS-encoding genes of the *hrp*/*hrc* region are located in a 26-kilobase pathogenicity island (Huang *et al.*, 1988; Alfano *et al.*, 2000; Oh *et al.*, 2007). Central to its expression is the alternative sigma factor HrpL (Xiao *et al.*, 1994). HrpL is a member of the extracytoplasmic function (ECF) family of sigma factors and activates expression of the T3SS-encoding genes and type III effector genes by recognizing a cis-regulatory element, the *hrp*-box (Fouts *et al.*, 2002; Innes *et al.*, 1993). Their expression is usually low or repressed when cells are grown in rich media but induced to high levels when grown *in planta* or *in vitro* in *hrp*-inducing media (Huynh *et al.*, 1989; Rahme *et al.*, 1992; Xiao *et al.*, 1992).

Type III effectors of phytopathogens are collectively necessary for the bacteria to cause disease. A T3SS-deficient mutant is incapable of causing disease on its normally compatible host (Lindgren *et al.*, 1986; Niepold *et al.*, 1985). Moreover, studies of transgenic plants expressing individual type III effectors provide strong evidence that their functions are to dampen host PTI (Hauck *et al.*, 2003; Nomura *et al.*, 2006; Fu *et al.*, 2007; Underwood *et al.*, 2007; Shan *et al.*, 2008; Xiang *et al.*, 2008).

Some type III effectors can also betray the presence of the infecting bacterium to the host. Plants encode disease resistance proteins (R) that can perceive a single or limited number of cognate type III effector proteins (DeYoung and Innes, 2006; Jones and Dangl, 2006). Perception triggers a defense response that can be viewed as amplified PTI. This effector-triggered immunity (ETI) is frequently associated with visual evidence of a programmed cell death response called a hypersensitive response (HR; Greenberg and Yao, 2004).

Germane to the work described herein are three ETI-elicitors, AvrRpt2, AvrRpm1, and HopQ1-1. AvrRpt2 and AvrRpm1 are type III effectors of *P. syringae* that perturb the *Arabidopsis* protein RIN4. The R proteins RPS2 and RPM1, respectively, perceive their corresponding modifications to RIN4 to elicit ETI (Mackey *et al.*, 2002; Axtell and Staskawicz, 2003; Mackey *et al.*, 2003). HopQ1-1 of *P. syringae* pv *tomato* DC3000 (*Pto*DC3000) elicits ETI in tobacco and is a negative host range determinant; its deletion

from the genome enables *Pto*DC3000 to grow significantly more than the wild-type strain in tobacco (Wei *et al.*, 2007).

Type III effectors are challenging to study for many reasons. There is a great diversity of effectors even between strains of the same species. A draft genome sequence of the pathogen *P. syringae* pv *tomato* T1 (*Pto*T1) was completed and compared to the completed genome sequence of *Pto*DC3000 (Almeida *et al.*, 2009). Both are pathogens of tomato and their inventories of genes had a high degree of conservation. In striking contrast, the number of homologous type III effector genes common to both strains was much lower. This diversity in type III effector collections is likely in response to the selective pressures host defenses impose on pathogens (Jones and Dangl, 2006).

Deletion or overexpression of type III effector genes in pathogens does not often result in changes in phenotype (Chang *et al.*, 2004). This reflects the observation that pathogens have collections of type III effectors with overlapping functions (Kvitko *et al.*, 2009). The overlap in functions of homologous and especially non-homologous type III effectors is thus a significant impediment in genetic-based studies of their functions.

P. fluorescens 55 carrying the pHIR11 cosmid or its derivatives, has been a workhorse heterologous delivery system and a valuable resource for studying type III effectors (Kim *et al.*, 2002; Jamir *et al.*, 2004; Schechter *et al.*, 2004; Fujikawa *et al.*, 2006). This cosmid clone has the entire T3SS-encoding region of *P. syringae* pv *syringae* 61 (*Psy*61) and also a chaperone and its cognate type III effector gene, *shcA* and *hopA1* (aka *hopPsyA*) as well as 1.6 kb of the type III effector gene *avrE1* (Huang *et al.*, 1988). HopA1 elicits effector-triggered immunity (ETI) in tobacco and the Ws-0 cultivar of *Arabidopsis* (Jamir *et al.*, 2004; Gassmann, 2005). The presence of *hopA1* complicates the characterization of defined sets of type III effectors and it has been disrupted via marker-exchange mutagenesis (Fouts *et al.*, 2003; Jamir *et al.*, 2004).

One limitation to pHIR11 or its variants is that bacterial cells do not stably maintain them in the absence of antibiotic selection. After 24 hours of growth *in planta*, only 50% of a derivative of *Pto*DC3000 still carried pCPP2071, a derivative of pHIR11

(Fouts *et al.*, 2003). By four days, only 4% of the cells still had the cosmid. We have observed ~90% loss from *P. fluorescens* grown overnight in culture in the absence of selection (Unrath and Chang, unpublished). This rapid loss of the T3SS-encoding locus potentially limits the utility of pHIR11 and its derivatives for studying type III effectors *in planta*.

To address the limitations in stability and potential conflict with antibiotic resistance markers, we have developed a new system to deliver individual or defined sets of type III effector proteins directly into host cells. We used recombineering to transfer the *hrp*/*hrc* region from *Psy*61 cloned in a cosmid vector, into a mini-*Tn5* Tc plasmid (de Lorenzo *et al.*, 1990; Court *et al.*, 2002; Jamir *et al.*, 2004; Oh *et al.*, 2007). Recombineering is catalyzed by bacteriophage-encoded systems that inhibit bacteria RecBCD nucleases from degrading double-stranded linear DNA fragments and can be used to precisely delete, insert, or alter a DNA sequence (Court *et al.*, 2002; Sharan *et al.*, 2009). We then stably integrated the *hrp*/*hrc* region directly into the genome of *P. fluorescens* Pf0-1 to ensure a robust delivery system (Silby *et al.*, 2009). We present our development of Effector to Host Analyzer (EtHAn) and validate its use for characterizing type III effectors.

RESULTS

Development of a stable type III effector delivery system

We used a variety of PCR methods and recombineering to clone the *hrp/hrc* cluster from cosmid clone pLN18 into a mini-*Tn5* Tc vector (Fig. 4.1; de Lorenzo *et al.*, 1990; Court *et al.*, 2002; Jamir *et al.*, 2004). The genes we cloned spanned *hrpK* to *hrpH* and we will refer to this as the T3SS-encoding region. Despite evidence that HrpH may be injected directly into host cells, we purposefully included *hrpH* because its product is necessary for efficient translocation of other type III effectors (Oh *et al.*, 2007). Our design avoided inclusion of genes encoding *shcA-hopA1* and included <300 bp of the *avrE* coding region. We further modified the mini-*Tn5* Tc by replacing the Tet^R gene with a

Kan^R gene flanked by site-specific recombinase recognition sequences (FRT) to facilitate excision of the antibiotic marker from the genome.

We used Tn5 to mediate the direct and stable integration of the T3SS-encoding region into the genome of the soil bacterium *P. fluorescens* Pf0-1. We selected this strain for two reasons. The genome sequence of Pf0-1 has been completed and searches have failed to confidently identify a T3SS-encoding region or candidate type III effector genes (Ma *et al.*, 2003; Grant *et al.*, 2006; Silby *et al.*, 2009). Secondly, Pf0-1 is a soil bacterium not adapted for survival within a plant and is therefore expected to elicit PTI and be incapable of dampening PTI or eliciting ETI when injected directly into plants (Compeau *et al.*, 1988).

We also used the site-specific recombinase FLP to excise the Kan^R gene from the genome of Pf0-1. Through these manipulations, we developed EtHAn (Effector to Host Analyzer), an unmarked, PTI-eliciting bacterium capable of delivering individual or defined sets of type III effector proteins directly into cells of plants.

P. fluorescens Pf0-1 and the modified EtHAn cannot grow *in planta*

To validate our selection of Pf0-1 as the recipient strain for the T3SS-encoding region, we first determined its growth behavior *in planta* (Fig. 4.2). We infiltrated 1.0 x 10^6 cfu/ml of wild-type Pf0-1 and EtHAn into leaves of *Arabidopsis* and enumerated their growth. The compatible pathogen *Pto*DC3000 grew extensively *in planta* and its T3SS mutant (D*hrcC*) failed to grow significantly. Pf0-1 and EtHAn failed to grow even to levels comparable to the D*hrcC* mutant. EtHAn expressing the T3SS appears to grow even less than Pf0-1, but we note that this is more likely a reflection of differences in infiltrated bacterial concentrations (at day 0, EtHAn had lower cfu than Pf0-1). We also enumerated growth of EtHAn following inoculation with higher concentrations of bacteria and had similar results to those presented (data not shown). These results validated our selection of Pf0-1 as the recipient for the T3SS-encoding region and show that the T3SS is, by itself, insufficient to confer virulence.

EtHAn carries a functional T3SS

To demonstrate that EtHAn encodes a functional T3SS, we tested its ability to deliver type III effectors *in planta*. Delivery is easily detected through elicitation of ETI where perception of a type III effector protein by a corresponding R protein leads to a rapid hypersensitive response (HR). We tested whether EtHAn carrying either *avrRpm1* or *avrRpt2* would elicit an HR in *Arabidopsis* Col-0. Their protein products are perceived by corresponding R proteins, RPM1 and RPS2, leading to a visible HR approximately 6 hpi and 20 hpi, respectively (Kunkel *et al.*, 1993; Yu *et al.*, 1993; Bisgrove *et al.*, 1994).

We infiltrated 1.0 x 10^8 cfu/ml of EtHAn carrying *avrRpm1* or *avrRpt2* into leaves of *Arabidopsis* (Fig. 4.3a). At 6 hpi, the majority of leaves infected with *Pto*DC3000 or EtHAn carrying *avrRpm1*, showed visible HR phenotypes. By 20 hpi, the majority of the leaves infected with *Pto*DC3000 or EtHAn carrying *avrRpt2* had collapsed. By 28 hpi, with the exception of EtHAn carrying an empty vector, all infected leaves had collapsed. This latter observation emphasizes one additional advantage of using EtHAn as opposed to wild-type *Pto*DC3000 for studying ETI. Because Pf0-1 is not adapted for survival within a plant, it is unable to elicit any visible symptoms in *Arabidopsis*, unlike wild-type *Pto*DC3000, which causes massive tissue collapse because of disease.

We next asked whether EtHAn carrying a single type III effector gene would gain virulence. The type III effector AvrPto has global effects on perturbing host PTI because it interacts with PAMP receptors to dampen plant defense responses (Shan *et al.*, 2008; Xiang *et al.*, 2008). Transgenic plants overexpressing AvrPto support more growth of the D*hrcC* mutant of *Pto*DC3000 (Hauck *et al.*, 2003). These transgenics are also severely compromised in responding to PAMPs as assayed by enumerating callose deposition (Hauck *et al.*, 2003). Finally, analysis of microarray data for host transcriptional changes suggested that the changes caused by AvrPto account for the majority of changes caused by wild-type *Pto*DC3000 delivering its entire set of type III effectors (Hauck *et al.*, 2003). We therefore determined if expression of AvrPto in EtHAn would be sufficient to confer virulence to a soil bacterium (Fig. 4.2). EtHAn carrying *avrPto* did not have any significant increase in growth relative to EtHAn carrying an empty vector. Thus, our growth

enumeration studies not surprisingly indicated that a single type III effector protein was not sufficient to confer virulence to a soil bacterium.

EtHAn expresses the T3SS and type III effectors to high levels

Our design of EtHAn relied on HrpL-controlled expression of the T3SS and its type III effector genes. Because proteins encoded by Pf0-1 are necessary to activate proteins upstream of HrpL, there is potential that EtHAn will not express type III effector genes to native levels as in *P. syringae*. We used quantitative real-time PCR (qRT-PCR) to measure the relative expression levels of a T3SS-encoding gene, *hrcV* and the type III effector gene *avrPto*. To ensure that expression levels were reflective of transcriptional regulation and not copy number of the type III effector gene, we used Tn7 to integrate a single copy of full-length *avrPto* into an intergenic region of the genome of Pf0-1 (Chang *et al.*, 2005; Peters and Craig, 2001).

As expected, *hrcV* and *avrPto* were expressed to high levels in *Pto*DC3000 seven hours after shift to *hrp*-inducing media but were not significantly expressed in the D*hrpL* mutant (Fig. 4.4a). EtHAn and EtHAn + *avrPto* expressed *hrcV* and *avrPto* in the latter strain, to higher levels than wild-type *Pto*DC3000 (Fig. 4.4b). Note that expression is presented as relative to *Pto*DC3000; i.e., we would have expected a normalized fold expression value of one if EtHAn expressed the HrpL-regulated genes to levels similar to *Pto*DC3000. By 24 hours after shift, expression of both genes in *Pto*DC3000 was still detectable and significant but far less than their levels at seven hours after shift to *hrp*-inducing media. In contrast, expression of both genes continued to increase in EtHAn and EtHAn + *avrPto*, respectively. We also noticed that expression of both genes was higher in cells grown in KB media than in *Pto*DC3000 (data not shown). Together, these results indicate that HrpL-regulated genes are not under the same negative control as they are in *P. syringae*. Regardless of this observation, results indicate that EtHAn expresses HrpL-regulated genes to sufficient levels as compared to *P. syringae*.

Type III effectors delivered by EtHAn are sufficient to dampen PTI

Next, we determined whether EtHAn is suitable for studying the virulence functions of type III effector proteins. Several type III effectors, including AvrPto and HopM1 have demonstrable roles in significantly suppressing the deposition of callose when expressed directly in transgenic plants (Hauck *et al.*, 2003; Nomura *et al.*, 2006; Fu *et al.*, 2007; Underwood *et al.*, 2007). We therefore infiltrated leaves of *Arabidopsis* with EtHAn carrying *avrPto* or *hopM1* and enumerated the number of callose deposits (Fig. 4.5). We also enumerated the number of callose deposits in leaves infiltrated with *Pto*DC3000, its *ΔhrcC* mutant, and EtHAn carrying an empty vector as controls. Representative leaf pictures are presented (Fig. 4.5b).

As expected, the *ΔhrcC* mutant that is incapable of delivering type III effectors was unable to suppress callose deposition whereas *Pto*DC3000 significantly suppressed callose deposition. Infiltration of leaves with EtHAn without any type III effector genes resulted in even more callose deposits than the *ΔhrcC* mutant. In contrast, EtHAn carrying *avrPto* or *hopM1* significantly and reproducibly suppressed the deposition of callose as compared to EtHAn carrying an empty vector. These results indicate that EtHAn by itself elicits PTI and that type III effectors delivered by EtHAn are sufficient to dampen PTI. Thus, EtHAn is a suitable system for studying the virulence functions of type III effector proteins *in planta*.

Delivery of type III effectors can be generalized

The possibility that Pf0-1 cannot deliver all *P. syringae* type III effectors is unlikely given the observation that *P. fluorescens* 55 carrying pLN18 was sufficient to deliver all tested type III effectors (Jamir *et al.*, 2004; Schechter *et al.*, 2004). Regardless, we tested approximately one half of the type III effectors from *Pto*DC3000 for delivery by EtHAn. The type III effectors were expressed as single copy genes integrated via Tn7, and from their native promoters as translational fusions to *Δ79AvrRpt2* (Guttman *et al.*, 2002; Chang *et al.*, 2005). The tested type III effectors AvrE1, AvrPto1, HopAF1, HopAB2, HopAM1-1, HopC1, HopE1, HopI1, HopP1, HopQ1-1, HopX1, HopY1, (ShcA1)-HopA1,

(SchF2)-HopF2, and (SchO1)-HopO1 were all sufficient for EtHAn to elicit an HR in *Arabidopsis* encoding the corresponding *R* gene, *Rps2*. For type III effectors expressed from operons, we also included their upstream genes (included in parentheses). We therefore concluded that EtHAn can deliver most if not all type III effectors of *P. syringae*.

EtHAn by itself elicits a defense response in species other than *Arabidopsis*

Because wild-type Pf0-1 does not encode any of its own type III effector genes, we reasoned that EtHAn has potential applications for studying effectors in hosts other than *Arabidopsis*. We assayed the effects of EtHAn in leaves of tomato and *N. tabacum*. EtHAn by itself elicited a response in leaves of tomato (data not shown) and a spotty, inconsistent response in leaves of *N. tabacum* (Fig. 4.3b). This response was specific to EtHAn; tomato or tobacco infiltrated with *P. fluorescens* Pf0-1 had no responses (data not shown). These results suggest the T3SS-encoding locus encodes a protein that may elicit a defense response in tomato or tobacco.

Nevertheless, we asked if EtHAn could deliver type III effectors into leaves of *N. tabacum*. *Pto*DC3000 is not compatible with tobacco because it delivers the perceived type III effector HopQ1-1 and elicits ETI (Wei *et al.*, 2007). We therefore mobilized *hopQ1-1* into EtHAn and infiltrated the strain into leaves of tobacco. After ~24 hpi, *Pto*DC3000 elicited a robust HR. The compatible pathogen *P. syringae* pv *tabaci* 11528, in contrast elicited strong tissue collapse due to disease whereas its corresponding T3SS mutant (*DhrcV*) did not elicit a phenotype (data not shown). EtHAn carrying *hopQ1-1* elicited a strong HR similar to *Pto*DC3000 and more robust as compared to EtHAn alone.

DISCUSSION

Our goal is to understand the mechanisms by which a single pathogen uses its entire collection of type III effector proteins to dampen host defenses. One difficulty is that pathogens can deliver more than thirty different type III effectors, of which many share overlapping functions (Kvitko *et al.*, 2009). The redundancy within a collection of type III effectors is likely a consequence of the need to ensure robustness so that loss of

single type III effectors will not compromise virulence. However, this redundancy is clearly a significant challenge that we must overcome in order to understand the contributions of each type III effector during host-association.

To address this hurdle, we developed a stable system for delivering type III effector proteins into host cells. We moved a defined fragment of ~26 kb spanning a region necessary to regulate and assemble the T3SS, into a modified mini-*Tn5* Tc vector. We used Tn5-mediated transposition to stably integrate the T3SS-encoding region directly into the genome of the soil bacterium, *P. fluorescens* Pf0-1 to make EtHAn. We show that when lacking type III effector genes, EtHAn elicits PAMP-triggered immunity and is non-pathogenic on *Arabidopsis*.

Potentially, EtHAn may be less efficient in delivering type III effector proteins into host cells than pathogenic *Pto*DC3000 as was the case for *P. fluorescens* 55 carrying pLN18 (Schechter *et al.*, 2004). We did observe variability in the robustness of the HR between different type III effector-Δ79AvrRpt2 fusions (data not shown). This however, is consistent with previous results using *Pto*DC3000 to deliver fusions with Δ79AvrRpt2 or *P. fluorescens* carrying pLN18 to deliver type III effector fusions to adenylate cyclase (Schechter *et al.*, 2004; Chang *et al.*, 2005). Furthermore, our reliance on native promoters may have contributed to variable expression. Nevertheless, when carrying type III effector genes, EtHAn expressed and translocated sufficient amounts of their proteins into host cells to elicit ETI or dampen the PTI-associated deposition of callose. Therefore, EtHAn is a sufficient elicitor of PAMP-triggered immunity that can subsequently be engineered to study the roles of delivered type III effectors in perturbing host defenses.

EtHAn offers several advantages over currently used methods for studying plant-pathogen interactions. Because EtHAn is decorated with molecular patterns and can be injected into the apoplastic space of plants, it can be used directly to study host defenses without the need to introduce additional PAMPs. We selected *P. fluorescens* Pf0-1 for modification because it appears to be devoid of any endogenous type III effector genes as determined by surveying its completed genome sequence for homologous DNA

sequences (Grant *et al.*, 2006). Further, Pf0-1 appears to lack most, if not all, necessary virulence factors required for growth *in planta*. Thus, any observed phenotypes can be attributed to the delivered type III effector protein of interest. Additionally, because the T3SS-encoding region is stably integrated and the type III effector genes can be as well, EtHAn can be used to characterize type III effector proteins *in planta* over the course of days.

EtHAn has applications beyond characterizing type III effector proteins of *P. syringae*. Fusions between oomycete effectors and the amino-terminal domain of AvrRps4 or AvrRpm1 are delivered by *Pto*DC3000 directly into plant cells via its T3SS (Sohn *et al.*, 2007; Rentel *et al.*, 2008). However, in most cases the virulence functions of the oomyete effector will be masked by the functions conferred by the collection of type III effectors normally delivered by *Pto*DC3000 (Sohn *et al.*, 2007). In contrast, since EtHAn is apparently devoid of virulence factors, the functions for effectors are more apt to be observed. EtHAn has been successfully used to deliver the *Hyaloperonospora parasitica* effectors ATR1 and ATR13 fused to AvrRpm1 directly into *Arabidopsis* (Brian Staskawicz, personal communication). Furthermore, *P. fluorescens* 55 carrying pHIR11 can deliver type III effectors from *Xanthomonas* spp (Fujikawa *et al.*, 2006). Though we did not explicitly reconfirm this finding we expect EtHAn should behave similarly. EtHAn therefore has potential applications for characterizing virulence functions of other bacterial as well as fungal and oomycete effectors.

EtHAn has strong potential for studying type III effector functions in *Arabidopsis*. EtHAn by itself does not elicit any symptoms in 88 different accessions of *Arabidopsis* (Qingli Lu, Marc Nishimura, and Jeffery Dangl, personal communication). Most notably, EtHAn does not elicit an HR in the Ws-0 accession, which is further evidence that our recombineering of the T3SS-encoding locus avoided inclusion of the type III effector gene *hopA1* (Gassmann, 2005).

In contrast, however, in its current form, EtHAn may have limited use in other plant species. Our results suggest that a factor encoded by the T3SS-encoding locus elicits a weaker non-host defense response in tomato or tobacco. Alternatively, but less

likely is the possibility that the phenotype elicited by EtHAn is a consequence of the *Tn*5-mediated integration event into the genome of Pf0-1. Regardless of the cause, our observations are not in agreement with published reports showing that *P. fluorescens* Pf-55 carrying pHIR11 elicits an HR in tobacco, while the same strain carrying pLN18 (disruption in *hopA1*), does not (Jamir *et al.*, 2004; Schechter *et al.*, 2004). Therefore, we speculate that EtHAn elicits a defense response in tobacco because the T3SS is stably integrated and expressed at higher levels than in pathogenic *P. syringae*.

Our results highlight the utility of recombineering for manipulating DNA fragments. This method has been used to delete genes directly from the genome of the plant pathogen, *Erwinia amylovora* but we have not had success with recombineering directly in *P. syringae* (Zhao *et al.*, 2009; Chang, unpublished). Nevertheless, we demonstrate that recombineering can be used to alter large and otherwise challenging to manipulate fragments of DNA in *E. coli*. Other potential applications include modifying large genomic clones or plasmids such as binary vectors (Rozwadowski *et al.*, 2008).

To request EtHAn, please visit our webpage at: http://changlab.cgrb.oregonstate.edu/.

EXPERIMENTAL PROCEDURES

Bacterial strains, plant lines and growth conditions

The bacterial strains used in this study are *P. syringae* pv *tomato* DC3000, its T3SS structural mutant (D*hrcC*; Yuan and He, 1996), its T3SS regulatory mutant, (D*hrpL*; Zwiesler-Vollick *et al.*, 2002), and *P. fluorescens* Pf0-1 (Silby *et al.*, 2009), *Escherichia coli* DH5a and HB101lpir. Pseudomonads were grown at 28°C in King's B (KB) liquid media with shaking or on KB agar plates. For *in vitro* induction of HrpL-regulated genes, Pseudomonads were grown overnight in KB media, washed, and resuspended at OD_{600} = 0.1 in *hrp*-inducing media and grown for 7 or 24 hours (Huynh et al., 1989). *E. coli* were grown at 37°C in Luriae-Bertani (LB) liquid media with shaking or on LB agar plates. Antibiotics were used at the final concentrations of: 25 mg/ml rifampicin, 30 mg/ml

kanamycin (100 mg/ml for Pf0-1), 30 mg/ml chloramphenicol, 5 mg/ml tetracycline (50 mg/ml for Pf0-1), 25 mg/ml gentamycin (100 mg/ml for Pf0-1) and 100 mg/ml ampicillin. Concentrations listed were for *P. syringae* and *E. coli* unless otherwise noted.

Arabidopsis thaliana accession Col-0 and *Nicotiana tabacum* were grown in a controlled-environment growth chamber (9 hrs of day at 22°C, 15 hrs of night at 20°C) for 5~6 weeks.

Plasmid constructions

Plasmids used were pBBR1-MCS1, -MCS2, and -MCS5 (Kovach *et al.*, 1994; Kovach *et al.*, 1995), pRK2013 (Figurski and Helinski, 1979), mini-*Tn5* Tc (de Lorenzo *et al.*, 1990), pKD4 and pKD46 (Datsenko and Wanner, 2000), pBH474 (House *et al.*, 2004), pME3280a (Zuber *et al.*, 2003), pUX-BF13 (Bao *et al.*, 1991), and pLN18 (Jamir *et al.*, 2004). Oligonucleotides used are listed in supplementary table 1. All restriction enzymes were purchased from New England Biolabs (NEB; Ipswich, MA).

We first used recombinant and sticky-end PCR to fuse together two 0.5 kb-sized fragments flanking the T3SS-encoding region carried on pLN18 (Jamir *et al.*, 2004). A 0.5 kb fragment immediately downstream of *hrpK* was amplified in two separate reactions using Pfu and primer pairs JHC124 + JHC125 or JHC148 + JHC150 (fragments 1 and 1’ of Figure 4.1, respectively). A 0.5 kb fragment immediately upstream of *hrpH* (aka ORF1 of Conserved Effector Locus; Alfano *et al.*, 2000) was amplified in two separate reactions with primer pairs JHC126 + JHC128 or JHC151 + JHC128 (fragments 2 and 2’, respectively). The products were gel-purified. Fragments 1 + 2 and 1’ + 2’ were mixed in approximately equal ratios, and each amplified in two separate reactions using Pfu and primer pairs JHC123 + JHC125 and JHC124 + JHC127, or JHC148 + JHC128 and JHC149 + JHC152, respectively. Recombined products carried an inserted unique *Xba*I or *Acc*65I site, respectively. Corresponding ~1.0 kb products were mixed in approximately equal ratios, incubated at 95°C for 5 minutes, and slowly cooled to room temperature leading to a proportion of the PCR products with overhanging ends compatible to recipient vectors.

Sticky-end products derived from the 1 + 2 fusion were gel-purified and cloned into pBBR1-MCS5 cleaved with *Eco*RI + *Xho*I and subsequently subcloned into pBBR1-MCS1 cleaved with *Sac*I + *Xho*I (Kovach *et al.*, 1994; Kovach *et al.*, 1995). Sticky-end products from the 1' + 2' fusion were gel-purified and cloned into mini-*Tn5* Tc cleaved with *Not*I and treated with calf-intestinal phosphatase (CIP).

Recombineering

We digested pBBR1-MCS1 carrying the 1 + 2 fusion with *Xba*I and transformed the gel-purified, linear fragment into electrocompetent cells made from arabinose-induced (10 mM) DH5a cells carrying pKD46 and pLN18 (de Lorenzo *et al.*, 1990; Jamir *et al.*, 2004). Transformants were selected on chloramphenicol. Successful transfer of the entire ~26 kb T3SS-encoding region was confirmed by restriction digestion and PCR with pairs of primers that amplified different fragments along the length of the T3SS-encoding region.

We next cleaved mini-*Tn5* Tc carrying the 1' + 2' fusion with *Acc*65I and transformed the gel-purified, linear fragment into electrocompetent cells made from arabinose-induced (10 mM) HB101lpir cells carrying pKD46 and pBBR1-MCS1 with the T3SS-encoding region. Transformants were selected on tetracycline. Successful transfer of the T3SS-encoding region was confirmed by restriction digestion and PCR with pairs of primers that amplified different fragments along the length of the T3SS-encoding region. We also used recombineering to replace the Tet^R gene of mini-*Tn5* Tc vector with Kan^R flanked by site-specific recombinase FRT sequences. We used a two-step PCR to amplify Kan^R flanked by FRT sites from pKD4 with CAT0001 and CAT0002 and subsequently with CAT0003 and CAT0004 to include 50 bp of homology to either side of the Tet^R gene of mini-*Tn5* Tc (de Lorenzo *et al.*, 1990). Recombineering was done as described above. Transformants were selected on kanamycin and successful recombineering was confirmed via PCR and lack of growth on LB + Tet.

Plasmid mobilization

Plasmids were mobilized into recipients as previously described (Chang *et al.*, 2005). For mobilizing *Tn*7-based vectors, cells were mixed at a ratio of 10:1:1:1 with the latter being *E. coli* carrying pUX-BF13 (Bao *et al.*, 1991). Integration of genes into the genome of *P. fluorescens* Pf0-1 was confirmed using PCR. Eviction of Kan^R was mediated by pBH474, which encodes the FLP site-specific recombinase and confirmed by replica plating on KB agar plates with and without kanamycin (House *et al.*, 2004). Finally, cells resistant to 5% sucrose were selected to identify those that lost pBH474.

Quantitative Real Time PCR (qRT-PCR)

Cells were collected and immediately suspended in RNAprotect (Qiagen, Valencia, CA), and stored at -80°C. RNA was extracted using RNeasy according to the instructions of the manufacturer (Qiagen, Valencia, CA), and subsequently treated with DNase I (NEB, Ipswich, MA). Qualities of each RNA preparations were assessed on 1.0 X FA, 1.2% agarose gels. We measured the quantity of RNA using a Nanodrop ND-1000 (Thermo Scientific, Wilmington, DE).

One mg of total RNA from each sample was used to synthesize single stranded cDNA according to the instructions of the manufacturer (Superscript III; Invitrogen, Carlsbad, CA). We used reverse transcriptase PCR to determine the quality of the cDNA. We used oligonucleotides 23S-T and 23S-B to amplify 23S rRNA from 1.0 ng of each sample. We also used hrpEF-T and hrpEF-B oligonucleotides in separate reactions to confirm the complete removal of DNA from each sample. These oligonucleotides span the intergenic region of separately transcribed genes *hrpE* and *hrpF* (Rahme *et al.*, 1991). One ng of single-stranded cDNA was used in SYBR-Green qRT-PCR on a CFX96 real-time machine (BioRad, Hercules, CA). Oligonucleotides hrcV-T and hrcV-B as well as avrPto-T and avrPto-B were used to amplify portions of *hrcV* and *avrPto*, respectively. All reactions were done in triplicate and normalized to 23S rRNA. To calculate efficiencies for each primer pair, we did qRT-PCR on 1.0 ng, 0.1 ng, and 0.01 ng of single-stranded cDNA made

from *Pto*DC3000 grown in *hrp*-inducing media for 24 hrs. Efficiency-corrected $2^{-\Delta\Delta Ct}$ values were determined using the CFX Manager software (Biorad; Pfaffl, 2001).

***In planta* assay**

Bacterial cells were grown overnight in KB with appropriate antibiotics, washed and resuspended in 10 mM $MgCl_2$. For *in planta* growth assay, we resuspended *P. syringae* and *P. fluorescens* to an OD_{600} of 0.002 (1.0 x 10^6 cfu/ml). We used 1.0 ml syringes lacking needles to infiltrate bacterial suspensions into the abaxial side of leaves of 5 ~ 6 week-old plants. We used a number 2 cork borer to core four discs for each triplicate of each treatment at 0 and 3 days post inoculation (dpi). Leaf discs were ground to homogeneity in 10 mM $MgCl_2$, serially diluted, and plated on KB with appropriate antibiotics. Colonies were enumerated once visible. Experiments were repeated at least three times. For HR assays, cells were infiltrated in leaves at an OD_{600} of 0.2 (equivalent to 1.0 x 10^8 cfu/ml). Phenotypes were scored starting at six hours post inoculation (hpi) and examined up until disease symptoms were visible.

Callose Staining

We collected leaves 7 hpi, and cleared them of chlorophyll by heating at 65°C in 70% EtOH. Leaves were then rinsed and stained in aniline blue staining solution (150 mM K_2HPO_4, 0.1% aniline blue, pH 8.4) overnight. Leaves were placed in 70% glycerol containing 1.0% of the aniline blue staining solution and mounted onto microscope slides. We viewed the leaves using a light microscope (BX40, Olympus) with a filter under ultraviolet 10X magnification. Digital images of 10 fields per leaf were taken (Magnafire, Optronics) and 15 leaves were imaged per treatment. We used a custom Perl script to enumerate callose deposits.

ACKNOWLEDGEMENTS

We thank Mark Silby and Stuart Levy for *P. fluorescens* Pf0-1, Brenda Schroeder and Michael Kahn for pBH454, Dieter Haas for pME3280a and pUX-BF13, Sheng Yang He

for Δ*hrcC*, Brian Staskawicz for *ΔhrpL*, and Alan Collmer for pLN18 and the *ΔhrcV* mutant of *Pta* 11528. We thank Jeffery Dangl, Jason Cumbie, Rebecca Pankow, and Philip Hillebrand for their assistance and critical reading of this manuscript. We gratefully acknowledge Jim Carrington for use of his light microscope. Finally, we thank Ethan S. Chang for his assistance and inspiring the naming of the type III effector delivery system. This research was supported in part by start-up funds from Oregon State University to JHC and a grant from the National Research Initiative of the USDA Cooperative State Research, Education, and Extension Service (Grant 2008-35600-18783).

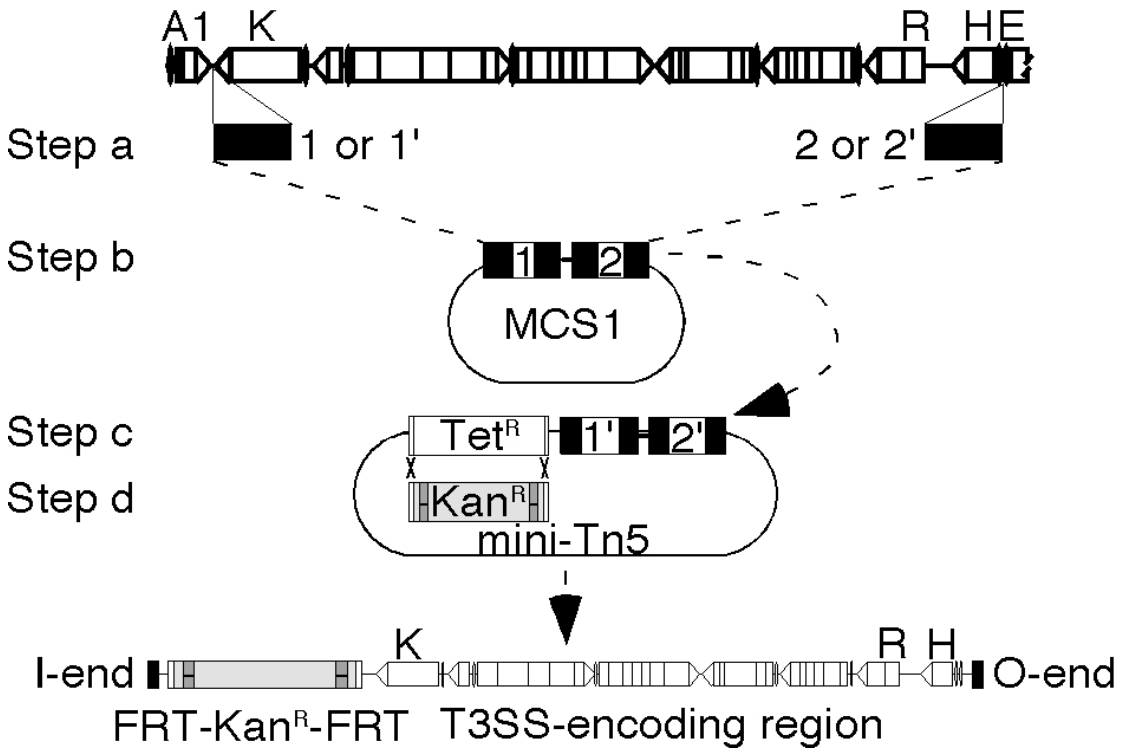


Figure 4.1. Construction of EtHAn.

The top bar represents the cloned insert of pLN18. Boxes represent protein-coding regions and triangles indicate their direction of transcription. Black vertical discs represent *hrp*-boxes. Some protein-coding regions are labeled for landmarks; abbreviations are: A1 = *shcA1-hopA1*, K = *hrpK*, R = *hrpR*, H = *hrpH*, and E = *avrE1*. The *avrE1* gene is truncated in pLN18 (broken line). Step a: Two 0.5 kb fragments (black boxes; 1, 1', 2 and 2') flanking the T3SS-encoding region were amplified and recombined. The 1 + 2 and 1' + 2' recombined products were cloned into pBBR1-MCS1 (MCS1) or mini-*Tn5* Tc, respectively. Steps b and c: Vectors containing recombined flanking DNA fragments were linearized by restriction digestion and used to sequentially capture the T3SS-encoding region by recombineering, first into pBBR1-MCS1 (step b) then to mini-*Tn5* Tc (step c). Step d: Recombineering was used to replace the Tet^R gene with the Kan^R gene flanked by FRT sites (gray bars with a horizontal dashed line). The resulting DNA fragment between the *Tn5*-repeats (I and O-ends) is shown at the bottom. The orientation of the T3SS-encoding region relative to the *Tn5* repeats is unknown and arbitrarily presented in one direction. The entire fragment was stably integrated into the genome of *P. fluorescens* Pf0-1 via *Tn5* transposition and the Kan^R gene was excised using the site-specific recombinase FLP (not shown).

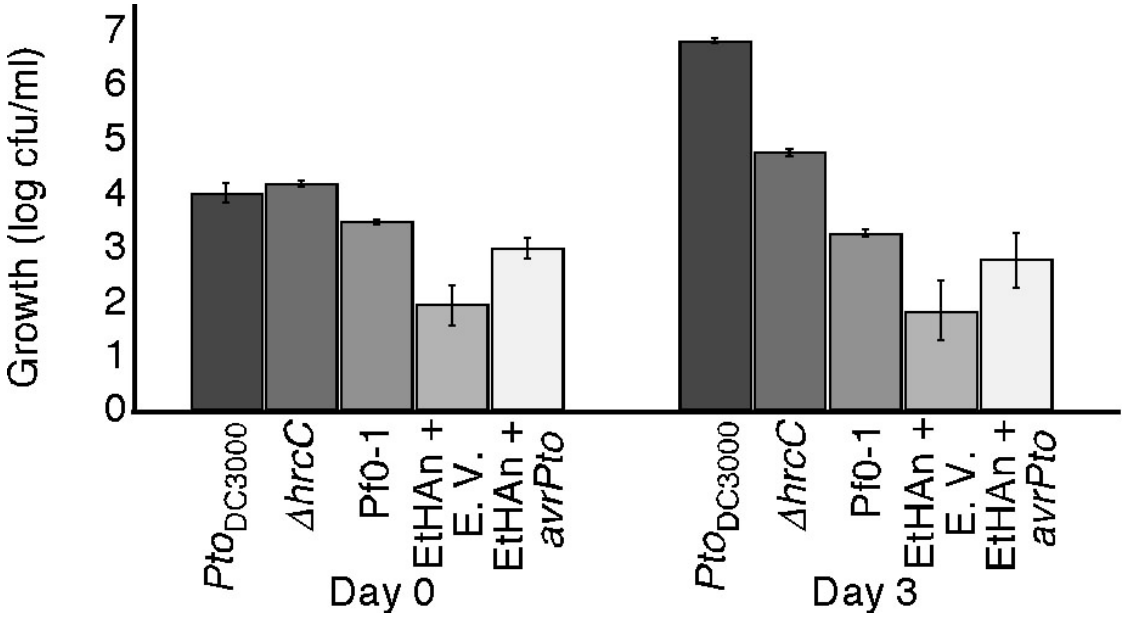


Figure 4.2. Enumeration of bacterial growth *in planta*.
Wild-type *Pto*DC3000, a D*hrcC* mutant, Pf0-1, EtHAn carrying an empty vector (E.V.), and EtHAn carrying *avrPto* were assessed for growth in *Arabidopsis* over a period of 3 dpi. Each time point was measured in triplicate; standard errors are presented. Experiments were repeated at least three times with similar results.

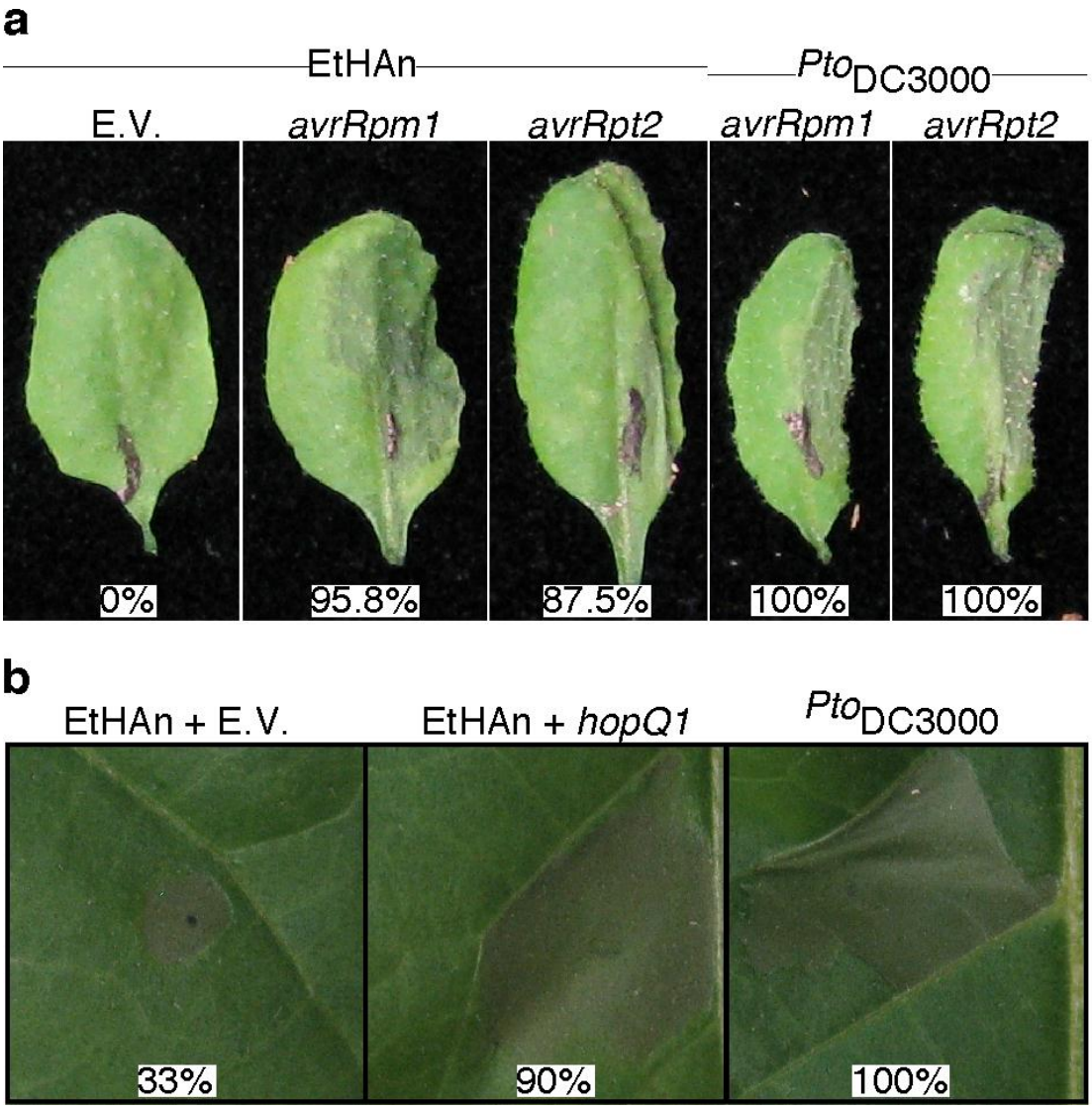


Figure 4.3. EtHAN has a functional type III secretion system.
a) EtHAn and *Pto*DC3000 carrying *avrRpm1* and *avrRpt2* elicited a strong HR in *Arabidopsis* Col-0. Twenty-four leaves were infiltrated for each strain and the percentages of responding leaves are presented at the bottom of each panel. Experiments were repeated at least three times with similar results. Black lines denote infiltrated leaves. b) EtHAn carrying full-length *hopQ1-1* elicited an HR in *N. tabacum*. EtHAn by itself elicited a spotty and inconsistent HR. Eighteen leaves were infiltrated for each strain and the percentages of responding leaves are shown. This experiment was repeated three times with similar results.

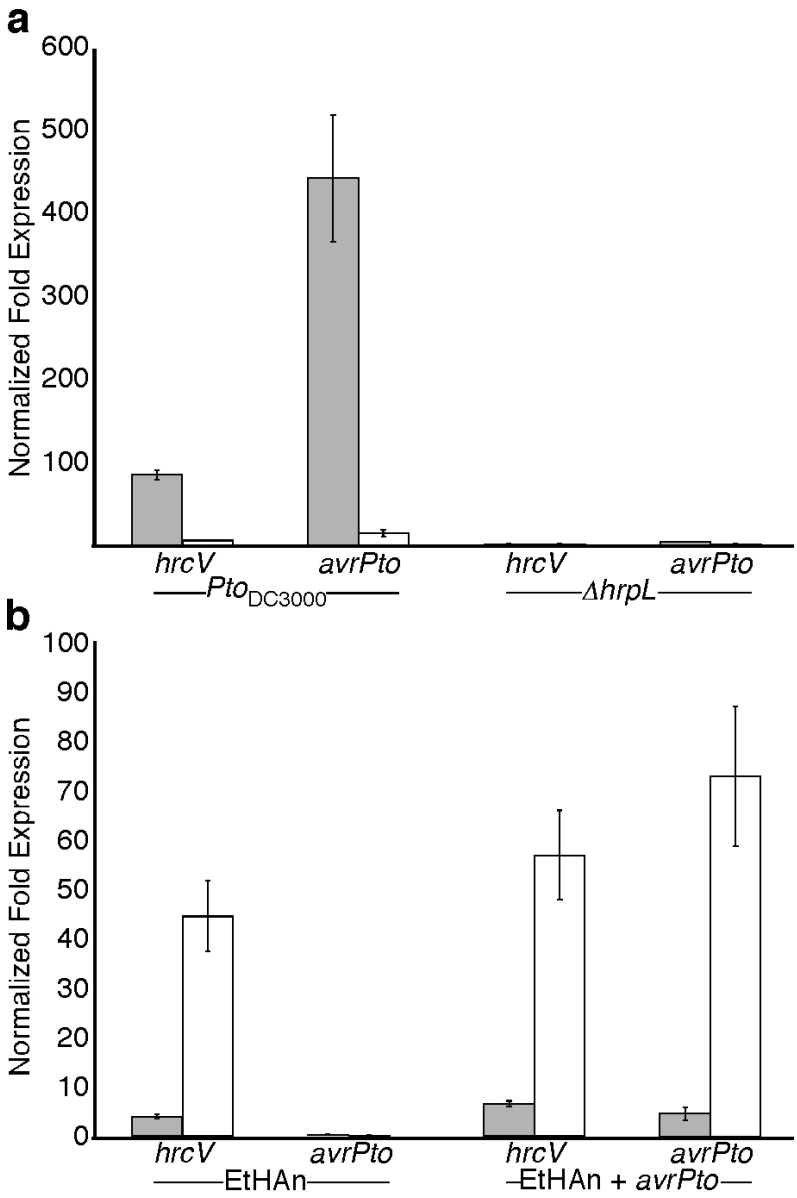


Figure 4.4. Expression of HrpL-regulated genes in EtHAn.
a) Normalized fold expression of *hrcV* and *avrPto* in *Pto*DC3000 and D*hrpL* were measured using the efficiency corrected 2^{-DDCt} qRT-PCR method. Gene expression was normalized to 23S rRNA and calculated relative to corresponding genes of cells grown in KB media. Expression was assessed at 7 (gray bars) and 24 (white bars) hours after shift to *hrp*-inducing media. b) Normalized fold expression of *hrcV* and *avrPto* in EtHAn or EtHAn carrying *avrPto* were measured at 7 (gray bars) and 24 (white bars) hours after shift to *hrp*-inducing media. Expression was normalized to 23S rRNA and calculated relative to corresponding genes in *Pto*DC3000 grown in *hrp*-inducing media (a). All genes were measured in triplicate. Standard error is shown.

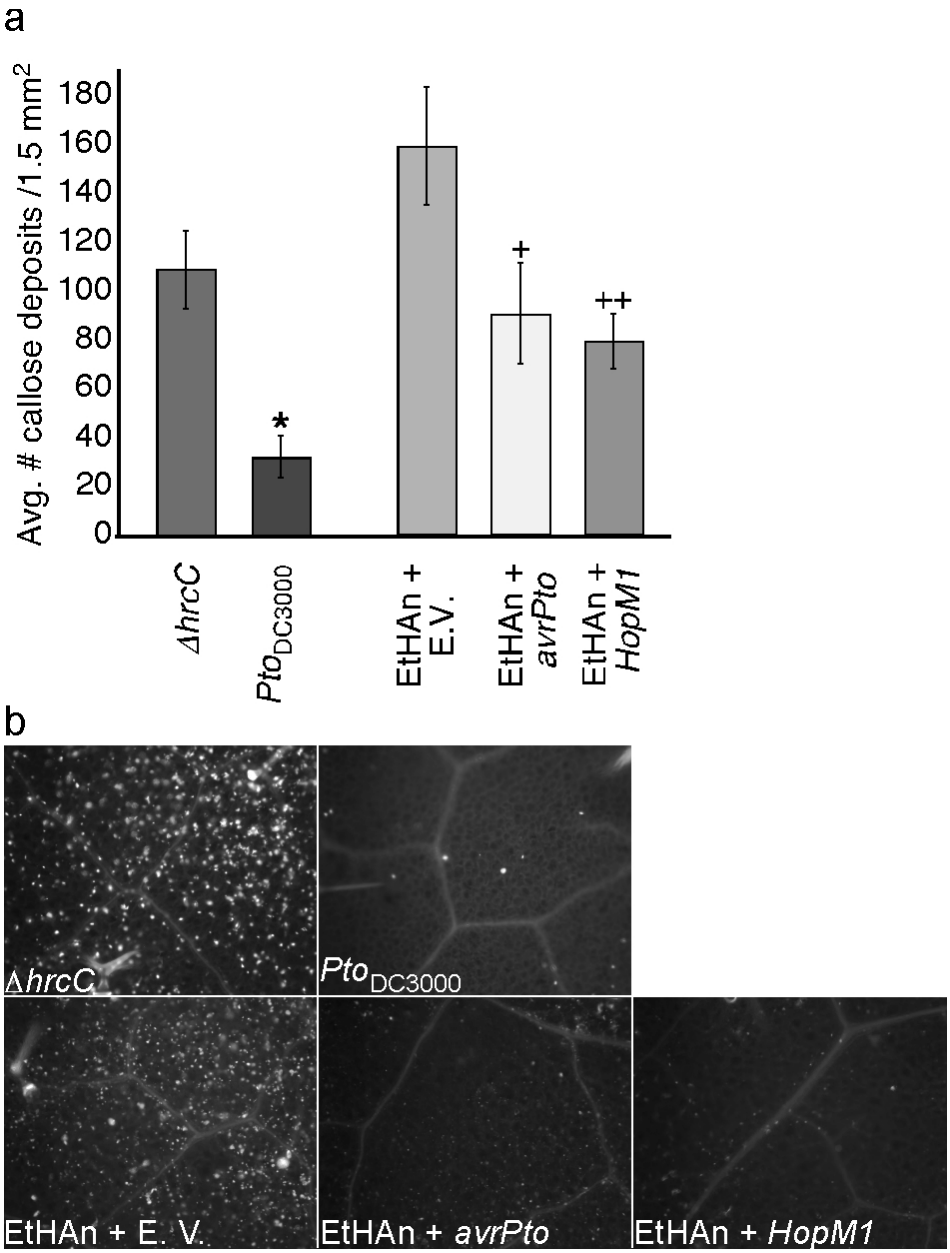


Figure 4.5. EtHAn carrying *avrRpm1* or *hopM1* dampens the callose response.
a) Average number of callose deposits per field of view (1.5 mm^2). Fifteen leaves were photographed per treatment with ten fields photographed per leaf. Significance of differences between averages were determined using Student's *t*-test (*Pto*DC3000 versus *hrcC* or EtHAn carrying *avrPto* or *hopM1* versus empty vector control,). $^{*,+}$ Denotes a p-value < 0.05; $^{++}$denotes a p-value < 0.01. This experiment was repeated multiple times with similar results. b) Representative gray-scaled images of *Arabidopsis* leaves stained with aniline blue for callose deposition.

Conclusions and Future Directions

William J. Thomas

The type III secretion system (T3SS) is used by many host-associated bacteria of agricultural importance to establish host-microbe relationships. The T3SS and its associated type III effector (T3E) proteins function in the suppression of immunity in host cells, a role traditionally characterized as "virulence". Given this association with virulence and its ubiquity among many devastating phytopathogenic bacteria, it is perhaps not surprising that decades of T3SS-related research were primarily pathogen-driven. However, mounting evidence of the prevalence of this mechanism in host-associated bacteria with other symbiotic lifestyles, especially in the critically important nitrogen-fixing rhizobia, has led to a re-conceptualization of the implications of T3E-mediated manipulations of immunity. In the case of rhizobia, T3Es are demonstrably necessary for efficient nodulation of the host legume, and loss of even one or two T3Es can change the host range of a microsymbiont. The importance of T3Es to the entire spectrum of host-bacteria symbioses underscores the need for characterization of this important class of proteins in mutualist as well as pathogenic bacteria.

The work presented in this thesis contributes to this process of characterization by significantly expanding the inventory of T3Es identified as having roles in mutualist interactions. The first study presented herein describes the development of a genome-enabled search for T3E-encoding genes specific to rhizobia and its application to elucidating the T3E repertoires of two species with T3SSs of ancient lineage. This work yielded the surprising finding that collections of T3Es were highly conserved between phylogenetically disparate strains within a species, and that there was even significant overlap in T3E repertoire content between the two species examined, *Sinorhizobium fredii* and *Bradyrhizobium japonicum.* These findings represent a departure from the "dogma" of the co-evolution between T3Es and plant immunity as modeled in phytopathogenic bacteria, which proposes an evolutionary arms race between host and microsymbiont. While there is considerable empirical evidence for this model for pathogenic bacteria, our findings indicate that the selective pressures involved in the co-evolution between collections of T3Es in bacteria and immunity in plants exist on a spectrum that mirrors the spectrum of symbiotic lifestyles. The T3E collections of

mutualist rhizobia experience purifying selection, resulting in static T3E repertoires in equilibrium with host immune proteins, representing an evolutionary “armistice”.

We also employed the same genome-enabled T3E search to *Mesorhizobium loti* MAFF303099. We selected this strain because based on current knowledge, it appears to be the sole strain of this species that relies on a T3SS to infect its host plant and is thus hypothesized to have recently acquired its T3SS locus. Our results generated a picture of a nascent T3E repertoire, assembled both from conserved T3Es found in other rhizobia that are associated with the T3SS and from T3Es unique to *Mesorhizobium loti* MAFF303099. The latter group is of particular interest, as it appears illustrative of the innovation of T3Es in the context of a newly-acquired T3SS locus, mediated by an unusually high concentration of insertion sequences in a recombination “hot spot”. In a particularly tantalizing example of T3E innovation, our analyses, when taken in the context of previous studies, indicate that at some point during or following the recombination event that resulted in the acquisition of its T3SS, a sequence of DNA in MAFF303099 symbiosis island that coded for a type IV effector protein homolog was inverted in such a way as to bring it under the regulation of a *tts*-box immediately upstream of the newly-acquired T3SS locus. This is made even more noteworthy by the fact that the homolog of this *tts*-box in R7A is the only one detected in the symbiosis island sequence of that organism. Furthermore, despite the apparent recent acquisition of the T3E collection in MAFF303099, several of these proteins are demonstrably necessary for efficient nodulation of host legumes (Sánchez *et al.*, 2012). Taken as a whole, the two rhizobia studies describe an evolutionary process beginning with the acquisition of the T3SS, the innovation and acquisition of T3Es to form a repertoire, and ending with a stable collection of T3Es in evolutionary equilibrium with host immunity.

In total, our studies added over 50 new families to the inventory of T3Es of rhizobia. This cataloging of T3Es is a necessary prelude to characterizing the role of T3Es in symbioses and how they shape the host range of microsymbionts. Future challenges include the elucidation of molecular functions and host targets of individual T3Es from these repertoires in order to dissect their respective contributions to the manipulations

of host cellular machinery for the benefit of the microsymbiont. With that in mind, we developed the effector protein delivery system described in this thesis, the Effector-to-Host Analyzer, or EtHAn. We used recombineering and traditional cloning methods to capture the entire 26-kb T3SS locus of a phytopathogenic bacterium and integrate it into the genome of a non-host-associated plant commensalist bacterium that lacked a native T3SS and T3Es. We demonstrated the utility of this system for interrogating phenotypes associated with individual T3Es of phytopathogenic bacteria, which future work can extend to the characterization of effector proteins with other symbiotic lifestyles. Indeed, this work has already been extended beyond its initial applications to T3Es of phytopathogenic bacteria to effector proteins of non-bacterial plant pathogens. Several studies have demonstrated the efficacy of EtHAn in characterizing effector proteins of the phytopathogenic oomycete *Hyaloperonospora parastica* (Krasileva *et al.*, 2011; Fabro *et al.*, 2011; Goritschnig *et al.*, 2012). While the γ-Proteobacterial EtHAn has not yet been shown to translocate T3Es of rhizobia, I propose that rhizobial T3Es could be delivered by this system if EtHAn is used in concert with the effector detector vector described in these studies (Krasileva *et al.*, 2011; Fabro *et al.*, 2011; Goritschnig *et al.*, 2012). In this manner, we may extend the significance of this work to the characterization of effectors from many types of microorganisms with a wide range of symbiotic lifestyles.

Bibliography

Alfano, J.R., Charkowski, A.O., Deng, W.-L., Badel, J.L., Petnicki-Ocwieja, T., van Dijk, K. and Collmer, A. (2000) 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. *Proceedings of the National Academy of Sciences*, **97**, 4856-4861.

Almeida, N.F., Yan, S., Lindeberg, M., Studholme, D.J., Schneider, D.J., Condon, B., Liu, H., Viana, C.J., Warren, A., Evans, C., Kemen, E., MacLean, D., Angot, A., Martin, G.B., Jones, J.D., Collmer, A., Setubal, J.C. and Vinatzer, B.A. (2008) A Draft Genome Sequence of Pseudomonas syringae pv. tomato T1 Reveals a Type III Effector Repertoire Significantly Divergent from That of Pseudomonas syringae pv. tomato DC3000. *Molecular Plant-Microbe Interactions*, **22**, 52-62.

Altschul, S.F., Madden, T.L., SchÃ¤ffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, **25**, 3389-3402.

An, C. and Mou, Z. (2011) Salicylic Acid and its Function in Plant ImmunityF. *Journal of Integrative Plant Biology*, **53**, 412-428.

Ausmees, N., Kobayashi, H., Deakin, W.J., Marie, C., Krishnan, H.B., Broughton, W.J. and Perret, X. (2004) Characterization of NopP, a Type III Secreted Effector of Rhizobium sp. Strain NGR234. *Journal of Bacteriology*, **186**, 4774-4780.

Ausubel, F.M. (2005) Are innate immune signaling pathways in plants and animals conserved? *Nat Immunol*, **6**, 973-979.

Axtell, M.J. and Staskawicz, B.J. (2003) Initiation of RPS2-Specified Disease Resistance in Arabidopsis Is Coupled to the AvrRpt2-Directed Elimination of RIN4. *Cell*, **112**, 369-377.

Baltrus, D.A., Nishimura, M.T., Romanchuk, A., Chang, J.H., Mukhtar, M.S., Cherkis, K., Roach, J., Grant, S.R., Jones, C.D. and Dangl, J.L. (2011) Dynamic Evolution of Pathogenicity Revealed by Sequencing and Comparative Genomics of 19 <italic>Pseudomonas syringae</italic> Isolates. *PLoS Pathog*, **7**, e1002132.

Bao, Y., Lies, D.P., Fu, H. and Roberts, G.P. (1991) An improved Tn7-based system for the single-copy insertion of cloned genes into chromosomes of gram-negative bacteria. *Gene*, **109**, 167-168.

Barash, I. and Manulis-Sasson, S. (2007) Virulence mechanisms and host specificity of gall-forming Pantoea agglomerans. *Trends in Microbiology*, **15**, 538-545.

Bellato, C.M., Balatti, P.A., Pueppke, S.G. and Krishnan, H.B. (1996) Proteins from cells of Rhizobium fredii bind to DNA sequences preceding *nol*X, a flavonoid-inducible nod gene that is not associated with a nod box. *Molecular Plant-Microbe Interactions*, **9**, 457.

Bellato, C., Krishnan, H.B., Cubo, T., Temprano, F. and Pueppke, S.G. (1997) The soybean cultivar specificity gene noIX is present, expressed in a nodD-dependent manner, and of symbiotic significance in cultivar-nonspecific strains of Rhizobium (Sinorhizobium) fredii. *Microbiology*, **143**, 1381-1388.

Bogdanove, A.J., Koebnik, R., Lu, H., Furutani, A., Angiuoli, S.V., Patil, P.B., Van Sluys, M.-A., Ryan, R.P., Meyer, D.F., Han, S.-W., Aparna, G., Rajaram, M., Delcher, A.L., Phillippy, A.M., Puiu, D., Schatz, M.C., Shumway, M., Sommer, D.D., Trapnell, C., Benahmed, F., Dimitrov, G., Madupu, R., Radune, D., Sullivan, S., Jha, G., Ishihara, H., Lee, S.-W., Pandey, A., Sharma, V., Sriariyanun, M., Szurek, B., Vera-Cruz, C.M., Dorman, K.S., Ronald, P.C., Verdier, V.r., Dow, J.M., Sonti, R.V., Tsuge, S., Brendel, V.P., Rabinowicz, P.D., Leach, J.E., White, F.F. and Salzberg, S.L. Two New Complete Genome Sequences Offer Insight into Host and Tissue Specificity of Plant Pathogenic Xanthomonas spp. *Journal of Bacteriology*, **193**, 5450-5464.

Brzoska, P.M. and Signer, E.R. (1991) lpsZ, a lipopolysaccharide gene involved in symbiosis of Rhizobium meliloti. *Journal of Bacteriology*, **173**, 3235-3237.

Cai, R., Lewis, J., Yan, S., Liu, H., Clarke, C.R., Campanile, F., Almeida, N.F., Studholme, D.J., Lindeberg, M., Schneider, D., Zaccardelli, M., Setubal, J.C., Morales-Lizcano, N.P., Bernal, A., Coaker, G., Baker, C., Bender, C.L., Leman, S. and Vinatzer, B.A. (2011) The Plant Pathogen *Pseudomonas syringae* pv. *tomato* Is Genetically Monomorphic and under Strong Selection to Evade Tomato Immunity. *PLoS Pathog*, **7**, e1002130.

Cao, F., Yoshioka, K. and Desveaux, D. (2011) The roles of ABA in plant–pathogen interactions. *Journal of Plant Research*, **124**, 489-499.

Cascales, E. and Christie, P.J. (2003) The versatile bacterial type IV secretion systems. *Nat Rev Micro*, **1**, 137-149.

Chang, J.H., Goel, A.K., Grant, S.R. and Dangl, J.L. (2004) Wake of the flood: ascribing functions to the wave of type III effector proteins of phytopathogenic bacteria. *Current Opinion in Microbiology*, **7**, 11-18.

Chang, J.H., Urbach, J.M., Law, T.F., Arnold, L.W., Hu, A., Gombar, S., Grant, S.R., Ausubel, F.M. and Dangl, J.L. (2005) A high-throughput, near-saturating screen for type III effector genes from Pseudomonas syringae. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 2549-2554.

Chisholm, S.T., Coaker, G., Day, B. and Staskawicz, B.J. (2006) Host-Microbe Interactions: Shaping the Evolution of the Plant Immune Response. *Cell*, **124**, 803-814.

Compeau, G., Al-Achi, B., Platsouka, E. and Levy, S.B. (1988) Survival of rifampin-resistant mutants of Pseudomonas fluorescens and Pseudomonas putida in soil systems. *Applied and Environmental Microbiology*, **54**, 2432-2438.

Court, D.L., Sawitzke, J.A. and Thomason, L.C. (2002) Genetic engineering using homologous recombination. *Annual Review of Genetics*, **36**, 361-388.

Cui, H., Xiang, T. and Zhou, J.-M. (2009) Plant immunity: a lesson from pathogenic bacterial effector proteins. *Cellular Microbiology*, **11**, 1453-1461.

Cunnac, S.B., Lindeberg, M. and Collmer, A. (2009) Pseudomonas syringae type III secretion system effectors: repertoires in search of functions. *Current Opinion in Microbiology*, **12**, 53-60.

Cunnac, S.b., Chakravarthy, S., Kvitko, B.H., Russell, A.B., Martin, G.B. and Collmer, A. (2011) Genetic disassembly and combinatorial reassembly identify a minimal functional repertoire of type III effectors in Pseudomonas syringae. *Proceedings of the National Academy of Sciences*, **108**, 2975-2980.

Dai, W.-J., Zeng, Y., Xie, Z.-P. and Staehelin, C. (2008) Symbiosis-Promoting and Deleterious Effects of NopT, a Novel Type 3 Effector of Rhizobium sp. Strain NGR234. *Journal of Bacteriology*, **190**, 5101-5110.

Dale, C. and Welburn, S.C. (2001) The endosymbionts of tsetse flies: manipulating host-parasite interactions. *International Journal for Parasitology*, **31**, 628-631.

Dale, C., Plague, G.R., Wang, B., Ochman, H. and Moran, N.A. (2002) Type III secretion systems and the evolution of mutualistic endosymbiosis. *Proceedings of the National Academy of Sciences*, **99**, 12397-12402.

Dangl, J.L., Ritter, C.L., Gibbon, M.J., Mur, L.A.J., Wood, J.R., Goss, S., Mansfield, J.W., Taylor, J.D. and Vivian, A. (1992) Functional homologs of the Arabidopsis *RPM1* disease resistance gene in bean and pea. *The Plant Cell*, **4**, 1359-1369.

Dangl, J.L. and Jones, J.D.G. (2001) Plant pathogens and integrated defence responses to infection. *Nature*, **411**, 826-833.

Datsenko, K.A. and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences*, **97**, 6640-6645.

Dazzo, F.B., Truchet, G.L., Hollingsworth, R.I., Hrabak, E.M., Pankratz, H.S., Philip-Hollingsworth, S., Salzwedel, J.L., Chapman, K., Appenzeller, L. and Squartini, A. (1991) Rhizobium lipopolysaccharide modulates infection thread development in white clover root hairs. *Journal of Bacteriology*, **173**, 5371-5384.

De Hoff, P.L., Brill, L.M. and Hirsch, A.M. (2009) Plant lectins: the ties that bind in root symbiosis and plant defense. *Molecular Genetics and Genomics*, **282**, 1-15.

de Lorenzo, V., Herrero, M., Jakubzik, U. and Timmis, K.N. (1990) Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. *Journal of Bacteriology*, **172**, 6568-6572.

de Lyra, M.d.C.C.P., López-Baena, F.J., Madinabeitia, N., Vinardell, J.M., Espuny, M.d.R., Cubo, M.T., Bellogín, R.A., Ruiz-Sainz, J.E. and Ollero, F.J. (2006) Inactivation of the *Sinorhizobium fredii* HH103 *rhcJ* gene abolishes nodulation outer proteins (Nops) secretion and decreases the symbiotic capacity with soybean. *International Microbiology*, **9**, 125-133.

Deakin, W.J., Marie, C., Saad, M.M., Krishnan, H.B. and Broughton, W.J. (2005) NopA Is Associated with Cell Surface Appendages Produced by the Type III Secretion System of Rhizobium sp. Strain NGR234. *Molecular Plant-Microbe Interactions*, **18**, 499-507.

Deakin, W.J. and Broughton, W.J. (2009) Symbiotic use of pathogenic strategies: rhizobial protein secretion systems. *Nat Rev Micro*, **7**, 312-320.

Debener, T., Lehnackers, H., Arnold, M. and Dangl, J.L. (1991) Identification and molecular mapping of a single Arabidopsis thaliana locus determining resistance to a phytopathogenic Pseudomonas syringae isolate. *The Plant Journal*, **1**, 289-302.

Delcher, A.L., Bratke, K.A., Powers, E.C. and Salzberg, S.L. (2007) Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics*, **23**, 673-679.

Deng, X., Xiao, Y., Lan, L., Zhou, J.-M. and Tang, X. (2009) Pseudomonas syringae pv. phaseolicola Mutants Compromised for Type III Secretion System Gene Induction. *Molecular Plant-Microbe Interactions*, **22**, 964-976.

DeYoung, B.J. and Innes, R.W. (2006) Plant NBS-LRR proteins in pathogen sensing and host defense. *Nat Immunol*, **7**, 1243-1249.

Douglas, A.E. (2008) Conflict, Cheats and the Persistence of Symbioses. *New Phytologist*, **177**, 849-858.

Eddy, S.R. (1998) Profile hidden Markov models. *Bioinformatics*, **14**, 755-763.

Fabro, G., Steinbrenner, J., Coates, M., Ishaque, N., Baxter, L., Studholme, D.J., KÃ¶rner, E., Allen, R.L., Piquerez, S.J.M., Rougon-Cardoso, A., Greenshields, D., Lei, R., Badel, J.L., Caillaud, M.-C., Sohn, K.-H., Van den Ackerveken, G., Parker, J.E., Beynon, J. and Jones, J.D.G. (2011) Multiple Candidate Effectors from the Oomycete Pathogen <italic>Hyaloperonospora arabidopsidis</italic> Suppress Host Plant Immunity. *PLoS Pathog*, **7**, e1002348.

Fauvart, M. and Michiels, J. (2008) Rhizobial secreted proteins as determinants of host specificity in the rhizobium–legume symbiosis. *FEMS Microbiology Letters*, **285**, 1-9.

Figurski, D.H. and Helinski, D.R. (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *Proceedings of the National Academy of Sciences*, **76**, 1648-1652.

Fouts, D.E., Badel, J.L., Ramos, A.R., Rapp, R.A. and Collmer, A. (2003) A Pseudomonas syringae pv. tomato DC3000 Hrp (Type III Secretion) Deletion Mutant Expressing the Hrp System of Bean Pathogen P. syringae pv. syringae 61 Retains Normal Host Specificity for Tomato. *Molecular Plant-Microbe Interactions*, **16**, 43-52.

Freiberg, C., Fellay, R., Bairoch, A., Broughton, W.J., Rosenthal, A. and Perret, X. (1997) Molecular basis of symbiosis between Rhizobium and legumes. *Nature*, **387**, 394-401.

Fu, Z.Q., Guo, M., Jeong, B.-r., Tian, F., Elthon, T.E., Cerny, R.L., Staiger, D. and Alfano, J.R. (2007) A type III effector ADP-ribosylates RNA-binding proteins and quells plant immunity. *Nature*, **447**, 284-288.

Fujikawa, T., Ishihara, H., Leach, J.E. and Tsuyumu, S. (2006) Suppression of Defense Response in Plants by the avrBs3/pthA Gene Family of Xanthomonas spp. *Molecular Plant-Microbe Interactions*, **19**, 342-349.

Gage, D.J. and Margolin, W. (2000) Hanging by a thread: invasion of legume plants by rhizobia. *Current Opinion in Microbiology*, **3**, 613-617.

Galan, J.E. and Wolf-Watz, H. (2006) Protein delivery into eukaryotic cells by type III secretion machines. *Nature*, **444**, 567-573.

Gassmann, W. (2005) Natural Variation in the Arabidopsis Response to the Avirulence Gene hopPsyA Uncouples the Hypersensitive Response from Disease Resistance. *Molecular Plant-Microbe Interactions*, **18**, 1054-1060.

Goritschnig, S., Krasileva, K.V., Dahlbeck, D. and Staskawicz, B.J. (2012) Computational Prediction and Molecular Characterization of an Oomycete Effector and the Cognate Arabidopsis Resistance Gene. *PLoS Genet*, **8**, e1002502.

Gottfert, M., Rothlisberger, S., Kundig, C., Beck, C., Marty, R. and Hennecke, H. (2001) Potential Symbiosis-Specific Genes Uncovered by Sequencing a 410-Kilobase DNA Region of the Bradyrhizobium japonicum Chromosome. *Journal of Bacteriology*, **183**, 1405-1412.

Graham, M.A., Marek, L.F. and Shoemaker, R.C. (2002) PCR Sampling of disease resistance-like sequences from a disease resistance gene cluster in soybean. *Theoretical and Applied Genetics*, **105**, 50-57.

Graham, M.A., Marek, L.F. and Shoemaker, R.C. (2002) Organization, Expression and Evolution of a Disease Resistance Gene Cluster in Soybean. *Genetics*, **162**, 1961-1977.

Grant, S.R., Fisher, E.J., Chang, J.H., Mole, B.M. and Dangl, J.L. (2006) Subterfuge and Manipulation: Type III Effector Proteins of Phytopathogenic Bacteria. *Annual Review of Microbiology*, **60**, 425-449.

Greenberg, J.T. and Yao, N. (2004) The role and regulation of programmed cell death in plant–pathogen interactions. *Cellular Microbiology*, **6**, 201-211.

Guttman, D.S. and Greenberg, J.T. (2001) Functional Analysis of the Type III Effectors AvrRpt2 and AvrRpm1 of Pseudomonas syringae with the Use of a Single-Copy Genomic Integration System. *Molecular Plant-Microbe Interactions*, **14**, 145-155.

Guttman, D.S., Vinatzer, B.A., Sarkar, S.F., Ranall, M.V., Kettler, G. and Greenberg, J.T. (2002) A Functional Screen for the Type III (Hrp) Secretome of the Plant Pathogen Pseudomonas syringae. *Science*, **295**, 1722-1726.

Haapalainen, M., van Gestel, K., Pirhonen, M. and Taira, S. (2009) Soluble Plant Cell Signals Induce the Expression of the Type III Secretion System of Pseudomonas syringae and Upregulate the Production of Pilus Protein HrpA. *Molecular Plant-Microbe Interactions*, **22**, 282-290.

Haney, C.H., Riely, B.K., Tricoli, D.M., Cook, D.R., Ehrhardt, D.W. and Long, S.R. (2011) Symbiotic Rhizobia Bacteria Trigger a Change in Localization and Dynamics of the Medicago truncatula Receptor Kinase LYK3. *The Plant Cell Online*, **23**, 2774-2787.

Hauck, P., Thilmony, R. and He, S.Y. (2003) A Pseudomonas syringae type III effector suppresses cell wall-based extracellular defense in susceptible Arabidopsis plants. *Proceedings of the National Academy of Sciences*, **100**, 8577-8582.

He, S.Y., Nomura, K. and Whittam, T.S. (2004) Type III protein secretion mechanism in mammalian and plant pathogens. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, **1694**, 181-206.

Hempel, J., Zehner, S., GÃ¶ttfert, M. and Patschkowski, T. (2009) Analysis of the secretome of the soybean symbiont Bradyrhizobium japonicum. *Journal of Biotechnology*, **140**, 51-58.

Herre, E.A., Knowlton, N., Mueller, U.G. and Rehner, S.A. (1999) The evolution of mutualisms: exploring the paths between conflict and cooperation. *Trends in Ecology & Evolution*, **14**, 49-53.

Hidalgo, Ã., Margaret, I., Crespo-Rivas, J.C., Parada, M., Murdoch, P.d.S., Lopez, A., Buendia-Claveria, A.M., Moreno, J., Albareda, M., Gil-Serrano, A.M., Rodriguez-Carvajal, M.A., Palacios, J.M., Ruiz-Sainz, J.E. and Vinardell, J.M. (2010) The rkpU gene of Sinorhizobium fredii HH103 is required for bacterial K-antigen polysaccharide production and for efficient nodulation with soybean but not with cowpea. *Microbiology*, **156**, 3398-3411.

Hirsch, A.M., Lum, M.R. and Downie, J.A. (2001) What Makes the Rhizobia-Legume Symbiosis So Special? *Plant Physiology*, **127**, 1484-1492.

House, B.L., Mortimer, M.W. and Kahn, M.L. (2004) New Recombination Methods for Sinorhizobium meliloti Genetics. *Applied and Environmental Microbiology*, **70**, 2806-2815.

Huang, H.C., Schuurink, R., Denny, T.P., Atkinson, M.M., Baker, C.J., Yucel, I., Hutcheson, S.W. and Collmer, A. (1988) Molecular cloning of a Pseudomonas syringae pv. syringae gene cluster that enables Pseudomonas fluorescens to elicit the hypersensitive response in tobacco plants. *Journal of Bacteriology*, **170**, 4748-4756.

Hubber, A., Vergunst, A.C., Sullivan, J.T., Hooykaas, P.J.J. and Ronson, C.W. (2004) Symbiotic phenotypes and translocated effector proteins of the Mesorhizobium loti strain R7A VirB/D4 type IV secretion system. *Molecular Microbiology*, **54**, 561-574.

Huynh, T.V., Dahlbeck, D. and Staskawicz, B.J. (1989) Bacterial Blight of Soybean: Regulation of a Pathogen Gene Determining Host Cultivar Specificity. *Science*, **245**, 1374-1377.

Jackson, R.W., Vinatzer, B., Arnold, D.L., Dorus, S. and Murillo, J.s. (2011) The influence of the accessory genome on bacterial pathogen evolution. *Mobile Genetic Elements*, **1**, 55-65.

Jamir, Y., Guo, M., Oh, H.-S., Petnicki-Ocwieja, T., Chen, S., Tang, X., Dickman, M.B., Collmer, A. and R. Alfano, J. (2004) Identification of Pseudomonas syringae type III effectors that can suppress programmed cell death in plants and yeast. *The Plant Journal*, **37**, 554-565.

Jiang, G. and Krishnan, H.B. (2000) Sinorhizobium fredii USDA257, a Cultivar-Specific Soybean Symbiont, Carries Two Copies of y4yA and y4yB, Two Open Reading Frames That Are Located in a Region That Encodes the Type III Protein Secretion System. *Molecular Plant-Microbe Interactions*, **13**, 1010-1014.

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

Jones, K.M., Kobayashi, H., Davies, B.W., Taga, M.E. and Walker, G.C. (2007) How rhizobial symbionts invade plants: the Sinorhizobium-Medicago model. *Nat Rev Micro*, **5**, 619-633.

Joshi, R.K. and Nayak, S. (2011) Functional characterization and signal transduction ability of nucleotide-binding site-leucine-rich repeat resistance genes in plants. *Genetics and Molecular Research*, **10**, 2637-2652.

Kambara, K., Ardissone, S., Kobayashi, H., Saad, M.M., Schumpp, O., Broughton, W.J. and Deakin, W.J. (2009) Rhizobia utilize pathogen-like effector proteins during symbiosis. *Molecular Microbiology*, **71**, 92-106.

Kanazin, V., Marek, L.F. and Shoemaker, R.C. (1996) Resistance gene analogs are conserved and clustered in soybean. *Proceedings of the National Academy of Sciences*, **93**, 11746-11750.

Kaneko, T., Nakamura, Y., Sato, S., Asamizu, E., Kato, T., Sasamoto, S., Watanabe, A., Idesawa, K., Ishikawa, A., Kawashima, K., Kimura, T., Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Mochizuki, Y., Nakayama, S., Nakazaki, N., Shimpo, S., Sugimoto, M., Takeuchi, C., Yamada, M. and Tabata, S. (2000) Complete Genome Structure of the Nitrogen-fixing Symbiotic Bacterium Mesorhizobium loti (Supplement). *DNA Research*, **7**, 381-406.

Kaneko, T., Nakamura, Y., Sato, S., Minamisawa, K., Uchiumi, T., Sasamoto, S., Watanabe, A., Idesawa, K., Iriguchi, M., Kawashima, K., Kohara, M., Matsumoto, M., Shimpo, S., Tsuruoka, H., Wada, T., Yamada, M. and Tabata, S. (2002) Complete Genomic Sequence of Nitrogen-fixing Symbiotic Bacterium Bradyrhizobium japonicum USDA110. *DNA Research*, **9**, 189-197.

Kaneko, T., Maito, H., Hirakawa, H., Uchiike, N., Minamisawa, K., Watanabe, A. and Sato, S. (2011) Complete Genome Sequence of the Soybean Symbiont Bradyrhizobium japonicum Strain USDA6T. *Genes*.

Kereszt, A., Mergaert, P. and Kondorosi, E. (2011) Bacteroid Development in Legume Nodules: Evolution of Mutual Benefit or of Sacrificial Victims? *Molecular Plant-Microbe Interactions*, **24**, 1300-1309.

Kim, Y.J., Lin, N.-C. and Martin, G.B. (2002) Two Distinct Pseudomonas Effector Proteins Interact with the Pto Kinase and Activate Plant Immunity. *Cell*, **109**, 589-598.

Kimbrel, J.A., Givan, S.A., Halgren, A.B., Creason, A.L., Mills, D.I., Banowetz, G.M., Armstrong, D.J. and Chang, J.H. (2010) An improved, high-quality draft genome sequence of the Germination-Arrest Factor-producing Pseudomonas fluorescens WH6. *BMC Genomics*, **11**, 522.

Kimbrel, J.A., Givan, S.A., Temple, T.N., Johnson, K.B. and Chang, J.H. (2011) Genome sequencing and comparative analysis of the carrot bacterial blight pathogen, Xanthomonas hortorum pv. carotae M081, for insights into pathogenicity and applications in molecular diagnostics. *Molecular Plant Pathology*, **12**, 580-594.

Klassen, J.L. and Currie, C.R. (2012) Gene fragmentation in bacterial draft genomes: extent, consequences and mitigation. *BMC Genomics*, **10**, 14.

Kovach, M.E., Phillips, R.W., Elzer, P.H., Roop II, R.M. and Peterson, K.M. (1994) pBBR1MCS: a broad-host-range cloning vector. *Biotechniques*, **16**, 800-802.

Kovach, M.E., Elzer, P.H., Steven Hill, D., Robertson, G.T., Farris, M.A., Roop Ii, R.M. and Peterson, K.M. (1995) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene*, **166**, 175-176.

Krasileva, K.V., Zheng, C., Leonelli, L., Goritschnig, S., Dahlbeck, D. and Staskawicz, B.J. Global Analysis of *Arabidopsis* Downy Mildew Interactions Reveals Prevalence of Incomplete Resistance and Rapid Evolution of Pathogen Recognition. *PLoS ONE*, **6**, e28765.

Krause, A., Doerfel, A. and GÃ¶ttfert, M. (2002) Mutational and Transcriptional Analysis of the Type III Secretion System of Bradyrhizobium japonicum. *Molecular Plant-Microbe Interactions*, **15**, 1228-1235.

Krishnan, H.B., Lorio, J., Kim, W.S., Jiang, G., Kim, K.Y., DeBoer, M. and Pueppke, S.G. (2003) Extracellular Proteins Involved in Soybean Cultivar-Specific Nodulation Are Associated with Pilus-Like Surface Appendages and Exported by a Type III Protein Secretion System in Sinorhizobium fredii USDA257. *Molecular Plant-Microbe Interactions*, **16**, 617-625.

Krzywinski, M., Schein, J., Birol, Ä.n., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J. and Marra, M.A. (2009) Circos: An information aesthetic for comparative genomics. *Genome Research*, **19**, 1639-1645.

Kunkel, B.N., Bent, A.F., Dahlbeck, D., Innes, R.W. and Staskawicz, B. (1993) RPS2, an Arabidopsis disease resistance locus specifying recognition of Pseudomonas syringae strains expressing the avirulence gene avrRpt2. *The Plant Cell*, **5**, 865-875.

Kurtz, S., Phillippy, A., Delcher, A., Smoot, M., Shumway, M., Antonescu, C. and Salzberg, S. (2004) Versatile and open software for comparing large genomes. *Genome Biology*, **5**, R12.

Kvitko, B.H., Park, D.H., Velasquez, A., Wei, C.-F., Russell, A.B., Martin, G.B., Schneider, D.J. and Collmer, A. (2009) Deletions in the Repertoire of *Pseudomonas syringae* pv. *tomato* DC3000 Type III Secretion Effector Genes Reveal Functional Overlap among Effectors. *PLoS Pathog*, **5**, e1000388.

Lagesen, K., Hallin, P., Rodland, E.A., Staeˆrfeldt, H.-H., Rognes, T. and Ussery, D.W. (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Research*, **35**, 3100-3108.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. and Higgins, D.G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947-2948.

Law, R. and Lewis, D. (2008) Biotic environments and the maintenance of sexsome evidence from mutualistic symbioses. *Biological Journal of the Linnean Society*, **20**, 249-276.

Lewis, J.D., Guttman, D.S. and Desveaux, D. (2009) The targeting of plant cellular systems by injected type III effector proteins. *Seminars in Cell & Developmental Biology*, **20**, 1055-1063.

Li, W. and Godzik, A. (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, **22**, 1658-1659.

Lindeberg, M., Stavrinides, J., Chang, J.H., Alfano, J.R., Collmer, A., Dangl, J.L., Greenberg, J.T., Mansfield, J.W. and Guttman, D.S. (2005) Proposed Guidelines for a Unified Nomenclature and Phylogenetic Analysis of Type III Hop Effector Proteins in the Plant Pathogen Pseudomonas syringae. *Molecular Plant-Microbe Interactions*, **18**, 275-282.

Lindgren, P.B., Peet, R.C. and Panopoulos, N.J. (1986) Gene cluster of Pseudomonas syringae pv. "phaseolicola" controls pathogenicity of bean plants and hypersensitivity of nonhost plants. *Journal of Bacteriology*, **168**, 512-522.

Lopez-Baena, F.J., Vinardell, J.M., Perez-Montano, F., Crespo-Rivas, J.C., Bellogin, R.A., Espuny, M.d.R. and Ollero, F.J. (2008) Regulation and symbiotic significance of nodulation outer proteins secretion in Sinorhizobium fredii HH103. *Microbiology*, **154**, 1825-1836.

Lopez-Gomez, M., Sandal, N., Stougaard, J. and Boller, T. (2012) Interplay of flg22-induced defence responses and nodulation in Lotus japonicus. *Journal of Experimental Botany*, **63**, 393-401.

Lorio, J.C., Kim, W.S. and Krishnan, H.B. (2004) NopB, a Soybean Cultivar-Specificity Protein from Sinorhizobium fredii USDA257, Is a Type III Secreted Protein. *Molecular Plant-Microbe Interactions*, **17**, 1259-1268.

Lowe, T.M. and Eddy, S.R. (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research*, **25**, 955-964.

Luna, E., Pastor, V., Robert, J.r.m., Flors, V., Mauch-Mani, B. and Ton, J. (2011) Callose Deposition: A Multifaceted Plant Defense Response. *Molecular Plant-Microbe Interactions*, **24**, 183-193.

Ma, Q., Zhai, Y., Schneider, J.C., Ramseier, T.M. and Saier Jr, M.H. (2003) Protein secretion systems of Pseudomonas aeruginosa and P. fluorescens. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, **1611**, 223-233.

Ma, W., Dong, F.F.T., Stavrinides, J. and Guttman, D.S. (2006) Type III Effector Diversification via Both Pathoadaptation and Horizontal Transfer in Response to a Coevolutionary Arms Race. *PLoS Genet*, **2**, e209.

Ma, W. (2011) Roles of Ca2+ and cyclic nucleotide gated channel in plant innate immunity. *Plant Science*, **181**, 342-346.

Mackey, D., Holt, B.F., Wiig, A. and Dangl, J.L. (2002) RIN4 Interacts with Pseudomonas syringae Type III Effector Molecules and Is Required for RPM1-Mediated Resistance in Arabidopsis. *Cell*, **108**, 743-754.

Mackey, D., Belkhadir, Y., Alonso, J.M., Ecker, J.R. and Dangl, J.L. (2003) Arabidopsis RIN4 Is a Target of the Type III Virulence Effector AvrRpt2 and Modulates RPS2-Mediated Resistance. *Cell*, **112**, 379-389.

Mansfield, J.W. (2009) From bacterial avirulence genes to effector functions via the hrp delivery system: an overview of 25 years of progress in our understanding of plant innate immunity. *Molecular Plant Pathology*, **10**, 721-734.

Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Lu, F., Marchler, G.H., Mullokandov, M., Omelchenko, M.V., Robertson, C.L., Song, J.S., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., Zheng, C. and Bryant, S.H. CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Research*, **39**, D225-D229.

Margaret, I., Becker, A., Blom, J., Bonilla, I., Goesmann, A., GÃ¶ttfert, M., Lloret, J., Mittard-Runte, V., RÃ¼ckert, C., Ruiz-Sainz, J.E., Vinardell, J.M.a. and Weidner, S. Symbiotic properties and first analyses of the genomic sequence of the fast growing model strain Sinorhizobium fredii HH103 nodulating soybean. *Journal of Biotechnology*, **155**, 11-19.

Maria Lopez-Lara, I., Orgambide, G., Dazzo, F.B., Olivares, J. and Toro, N.s. (1995) Surface polysaccharide mutants of Rhizobium sp. (Acacia) strain GRH2: major requirement of lipopolysaccharide for successful invasion of Acacia nodules and host range determination. *Microbiology*, **141**, 573-581.

Marie, C., Broughton, W.J. and Deakin, W.J. (2001) Rhizobium type III secretion systems: legume charmers or alarmers? *Current Opinion in Plant Biology*, **4**, 336-342.

Marie, C., Deakin, W.J., Viprey, V., KopciÃ±ska, J., Golinowski, W., Krishnan, H.B., Perret, X. and Broughton, W.J. (2003) Characterization of Nops, Nodulation Outer Proteins, Secreted Via the Type III Secretion System of NGR234. *Molecular Plant-Microbe Interactions*, **16**, 743-751.

Marie, C., Deakin, W.J., Ojanen-Reuhs, T., Diallo, E., Reuhs, B., Broughton, W.J. and Perret, X. (2004) TtsI, a Key Regulator of Rhizobium Species NGR234 Is Required for Type III-Dependent Protein Secretion and Synthesis of Rhamnose-Rich Polysaccharides. *Molecular Plant-Microbe Interactions*, **17**, 958-966.

Mazurier, S., Lemunier, M., Hartmann, A., Siblot, S. and Lemanceau, P. (2006) Conservation of type III secretion system genes in Bradyrhizobium isolated from soybean. *FEMS Microbiology Letters*, **259**, 317-325.

McCann, H.C. and Guttman, D.S. (2008) Evolution of the type III secretion system and its effectors in plant-microbe interactions. *New Phytologist*, **177**, 33-47.

Meinhardt, L.W., Krishnan, H.B., Balatti, P.A. and Pueppke, S.G. (1993) Molecular cloning and characterization of a sym plasmid locus that regulates cultivar-specific nodulation of soybean by Rhizobium fredii USDA257. *Molecular Microbiology*, **9**, 17-29.

Moran, N.A., Degnan, P.H., Santos, S.R., Dunbar, H.E. and Ochman, H. (2005) The players in a mutualistic symbiosis: Insects, bacteria, viruses, and virulence genes. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 16919-16926.

Mudgett, M.B. and Staskawicz, B.J. (1999) Characterization of the Pseudomonas syringae pv. tomato AvrRpt2 protein: demonstration of secretion and processing during bacterial pathogenesis. *Molecular Microbiology*, **32**, 927-941.

Mudgett, M.B., Chesnokova, O., Dahlbeck, D., Clark, E.T., Rossier, O., Bonas, U. and Staskawicz, B.J. (2000) Molecular signals required for type III secretion and translocation of the Xanthomonas campestris AvrBs2 protein to pepper plants. *Proceedings of the National Academy of Sciences*, **97**, 13324-13329.

Najimi, M., Balado, M., Lemos, M.L. and Osorio, C.R. (2009) Genetic characterization of pAsa6, a new plasmid from Aeromonas salmonicida subsp. salmonicida that encodes a type III effector protein AopH homolog. *Plasmid*, **61**, 176-181.

Niepold, F., Anderson, D. and Mills, D. (1985) Cloning determinants of pathogenesis from Pseudomonas syringae pathovar syringae. *Proceedings of the National Academy of Sciences*, **82**, 406-410.

Nomura, K., DebRoy, S., Lee, Y.H., Pumplin, N., Jones, J. and He, S.Y. (2006) A Bacterial Virulence Protein Suppresses Host Innate Immunity to Cause Plant Disease. *Science*, **313**, 220-223.

Nürnberger, T., Brunner, F., Kemmerling, B. and Piater, L. (2004) Innate immunity in plants and animals: striking similarities and obvious differences. *Immunological Reviews*, **198**, 249-266.

Oh, H.-S. and Collmer, A. (2005) Basal resistance against bacteria in Nicotiana benthamiana leaves is accompanied by reduced vascular staining and suppressed by multiple Pseudomonas syringae type III secretion system effector proteins. *The Plant Journal*, **44**, 348-359.

Okazaki, S., Okabe, S., Higashi, M., Shimoda, Y., Sato, S., Tabata, S., Hashiguchi, M., Akashi, R., Gottfert, M. and Saeki, K. (2010) Identification and Functional Analysis of Type III Effector Proteins in Mesorhizobium loti. *Molecular Plant-Microbe Interactions*, **23**, 223-234.

Peck, M.C., Fisher, R.F. and Long, S.R. (2006) Diverse Flavonoids Stimulate NodD1 Binding to nod Gene Promoters in Sinorhizobium meliloti. *Journal of Bacteriology*, **188**, 5417-5427.

Perret, X., Staehelin, C. and Broughton, W.J. (2000) Molecular Basis of Symbiotic Promiscuity. *Microbiology and Molecular Biology Reviews*, **64**, 180-201.

Perret, X., Kobayashi, H. and Collado-Vides, J. (2003) Regulation of expression of symbiotic genes in Rhizobium sp. NGR234. *Indian Journal of Experimental Biology*, **41**, 1101-1113.

Peters, J.E. and Craig, N.L. (2001) Tn7: smarter than we thought. *Nat Rev Mol Cell Biol*, **2**, 806-814.

Petkau, A., Stuart-Edwards, M., Stothard, P. and Van Domselaar, G. (2010) Interactive microbial genome visualization with GView. *Bioinformatics*, **26**, 3125-3126.

Pfaffl, M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research*, **29**, e45.

Pitzschke, A., Schikora, A. and Hirt, H. (2009) MAPK cascade signalling networks in plant defence. *Current Opinion in Plant Biology*, **12**, 421-426.

Pop, M. and Salzberg, S.L. (2008) Bioinformatics challenges of new sequencing technology. *Trends in Genetics*, **24**, 142-149.

Postel, S. and Kemmerling, B. (2009) Plant systems for recognition of pathogen-associated molecular patterns. *Seminars in Cell & Developmental Biology*, **20**, 1025-1031.

Preston, G.M., Bertrand, N. and Rainey, P.B. (2001) Type III secretion in plant growth-promoting Pseudomonas fluorescens SBW25. *Molecular Microbiology*, **41**, 999-1014.

Pueppke, S.G. and Broughton, W.J. (1999) Rhizobium sp. Strain NGR234 and R. fredii USDA257 Share Exceptionally Broad, Nested Host Ranges. *Molecular Plant-Microbe Interactions*, **12**, 293-318.

Rafiqi, M., Bernoux, M., Ellis, J.G. and Dodds, P.N. (2009) In the trenches of plant pathogen recognition: Role of NB-LRR proteins. *Seminars in Cell & Developmental Biology*, **20**, 1017-1024.

Rahme, L.G., Mindrinos, M.N. and Panopoulos, N.J. (1992) Plant and environmental sensory signals control the expression of hrp genes in Pseudomonas syringae pv. phaseolicola. *Journal of Bacteriology*, **174**, 3499-3507.

Rainey, P.B. (1999) Adaptation of Pseudomonas fluorescens to the plant rhizosphere. *Environmental Microbiology*, **1**, 243-257.

Ramsay, J.P., Sullivan, J.T., Stuart, G.S., Lamont, I.L. and Ronson, C.W. (2006) Excision and transfer of the Mesorhizobium loti R7A symbiosis island requires an integrase IntS, a novel recombination directionality factor RdfS, and a putative relaxase RlxS. *Molecular Microbiology*, **62**, 723-734.

Rentel, M.C., Leonelli, L., Dahlbeck, D., Zhao, B. and Staskawicz, B.J. (2008) Recognition of the Hyaloperonospora parasitica effector ATR13 triggers resistance against oomycete, bacterial, and viral pathogens. *Proceedings of the National Academy of Sciences*, **105**, 1091-1096.

Rissman, A.I., Mau, B., Biehl, B.S., Darling, A.E., Glasner, J.D. and Perna, N.T. (2009) Reordering contigs of draft genomes using the Mauve Aligner. *Bioinformatics*, **25**, 2071-2073.

Robbertse, B., Yoder, R.J., Boyd, A., Reeves, J. and Spatafora, J.W. (2011) Hal: an Automated Pipeline for Phylogenetic Analyses of Genomic Data. *PLoS Currents*, **7**.

Robert-Seilaniantz, A., Grant, M. and Jones, J.D.G. (2011) Hormone Crosstalk in Plant Disease and Defense: More Than Just JASMONATE-SALICYLATE Antagonism. *Annual Review of Phytopathology*, **49**, 317-343.

Rodrigues, J.o.A., Lopez-Baena, F.J., Ollero, F.J., Vinardell, J.M.a., Espuny, M.a.d.R., BellogÃ-n, R.A., Ruiz-Sainz, J.E., Thomas, J.R., Sumpton, D., Ault, J. and Thomas-Oates, J. (2007) NopM and NopD Are Rhizobial Nodulation Outer Proteins: Identification Using LC-MALDI and LC-ESI with a Monolithic Capillary Column. *Journal of Proteome Research*, **6**, 1029-1037.

Rozwadowski, K., Yang, W. and Kagale, S. (2008) Homologous recombination-mediated cloning and manipulation of genomic DNA regions using Gateway and recombineering systems. *BMC Biotechnology*, **8**, 88.

Saad, M.M., Kobayashi, H., Marie, C., Brown, I.R., Mansfield, J.W., Broughton, W.J. and Deakin, W.J. (2005) NopB, a Type III Secreted Protein of Rhizobium sp. Strain NGR234, Is Associated with Pilus-Like Surface Appendages. *Journal of Bacteriology*, **187**, 1173-1181.

Sachs, Joel L., Mueller, Ulrich G., Wilcox, ThomasÂ P. and Bull, James J. (2004) The Evolution of Cooperation. *The Quarterly Review of Biology*, **79**, 135-160.

Sachs, J.L., Ehinger, M.O. and Simms, E.L. (2010) Origins of cheating and loss of symbiosis in wild Bradyrhizobium. *Journal of Evolutionary Biology*, **23**, 1075-1089.

Sachs, J.L., Essenberg, C.J. and Turcotte, M.M. (2011) New paradigms for the evolution of beneficial infections. *Trends in Ecology & Evolution*, **26**, 202-209.

Sachs, J.L., Russell, J.E. and Hollowell, A.C. (2011) Evolutionary Instability of Symbiotic Function in <italic>Bradyrhizobium japonicum</italic>. *PLoS ONE*, **6**, e26370.

Sachs, J.L., Skophammer, R.G. and Regus, J.U. (2011) Evolutionary transitions in bacterial symbiosis. *Proceedings of the National Academy of Sciences*, **108**, 10800-10807.

Sanabria, N.M., Huang, J.-C. and Dubery, I.A. (2010) Self/nonself perception in plants in innate immunity and defense. *Self Nonself.*, **1**, 40-54.

Sanchez, C., Iannino, F., Deakin, W.J., Ugalde, R.A. and Lepek, V.C. (2009) Characterization of the Mesorhizobium loti MAFF303099 Type-Three Protein Secretion System. *Molecular Plant-Microbe Interactions*, **22**, 519-528.

Sánchez, C., Mercante, V., Babuin, M.F. and Lepek, V.C. (2012) Dual effect of Mesorhizobium loti T3SS functionality on the symbiotic process. *FEMS Microbiology Letters*, **330**, 148-156.

Schechter, L.M., Roberts, K.A., Jamir, Y., Alfano, J.R. and Collmer, A. (2004) Pseudomonas syringae Type III Secretion System Targeting Signals and Novel Effectors Studied with a Cya Translocation Reporter. *Journal of Bacteriology*, **186**, 543-555.

Schechter, L.M., Guenther, J., Olcay, E.A., Jang, S. and Krishnan, H.B. (2010) Translocation of NopP by Sinorhizobium fredii USDA257 into Vigna unguiculata Root Nodules. *Applied and Environmental Microbiology*, **76**, 3758-3761.

Schmeisser, C., Liesegang, H., Krysciak, D., Bakkou, N., Le Quere, A., Wollherr, A., Heinemeyer, I., Morgenstern, B., Pommerening-RÃ¶ser, A., Flores, M., Palacios, R., Brenner, S., Gottschalk, G., Schmitz, R.A., Broughton, W.J., Perret, X., Strittmatter, A.W. and Streit, W.R. (2009) Rhizobium sp. Strain NGR234 Possesses a Remarkable Number of Secretion Systems. *Applied and Environmental Microbiology*, **75**, 4035-4045.

Scholze, H. and Boch, J. (2011) TAL effectors are remote controls for gene activation. *Current Opinion in Microbiology*, **14**, 47-53.

Schwessinger, B. and Zipfel, C. (2008) News from the frontline: recent insights into PAMP-triggered immunity in plants. *Current Opinion in Plant Biology*, **11**, 389-395.

Segonzac, C.c. and Zipfel, C. (2011) Activation of plant pattern-recognition receptors by bacteria. *Current Opinion in Microbiology*, **14**, 54-61.

Shan, L., He, P., Li, J., Heese, A., Peck, S.C., Nurnberger, T., Martin, G.B. and Sheen, J. (2008) Bacterial Effectors Target the Common Signaling Partner BAK1 to Disrupt Multiple MAMP Receptor-Signaling Complexes and Impede Plant Immunity. *Cell Host & Microbe*, **4**, 17-27.

Sharan, S.K., Thomason, L.C., Kuznetsov, S.G. and Court, D.L. (2009) Recombineering: a homologous recombination-based method of genetic engineering. *Nat. Protocols*, **4**, 206-223.

Shigenobu, S., Watanabe, H., Hattori, M., Sakaki, Y. and Ishikawa, H. (2000) Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. APS. *Nature*, **407**, 81-86.

Silby, M., Cerdeno-Tarraga, A., Vernikos, G., Giddens, S., Jackson, R., Preston, G., Zhang, X.-X., Moon, C., Gehrig, S., Godfrey, S., Knight, C., Malone, J., Robinson, Z., Spiers, A., Harris, S., Challis, G., Yaxley, A., Harris, D., Seeger, K., Murphy, L., Rutter, S., Squares, R., Quail, M., Saunders, E., Mavromatis, K., Brettin, T., Bentley, S., Hothersall, J., Stephens, E., Thomas, C., Parkhill, J., Levy, S., Rainey, P. and Thomson, N. (2009) Genomic and genetic analyses of diversity and plant interactions of Pseudomonas fluorescens. *Genome Biology*, **10**, R51.

Skorpil, P., Saad, M.M., Boukli, N.M., Kobayashi, H., Ares-Orpel, F., Broughton, W.J. and Deakin, W.J. (2005) NopP, a phosphorylated effector of Rhizobium sp. strain NGR234, is a major determinant of nodulation of the tropical legumes Flemingia congesta and Tephrosia vogelii. *Molecular Microbiology*, **57**, 1304-1317.

Sohn, K.H., Lei, R., Nemri, A. and Jones, J.D.G. (2007) The downy mildew effector proteins ATR1 and ATR13 promote disease susceptibility in Arabidopsis thaliana. *The Plant Cell*, **19**, 4077-4090.

Soto, M.J., Domínguez-Ferreras, A., Pérez-Mendoza, D., Sanjuán, J. and Olivares, J. (2009) Mutualism versus pathogenesis: the give-and-take in plant–bacteria interactions. *Cellular Microbiology*, **11**, 381-388.

Spaink, H.P. (2000) Root nodulation and infection factors produce by rhizobial bacteria. *Annual Review of Microbiology*, **54**, 257-288.

Stavrinides, J., McCann, H.C. and Guttman, D.S. (2008) Host–pathogen interplay and the evolution of bacterial effectors. *Cellular Microbiology*, **10**, 285-292.

Streit, W.R., Schmitz, R.A., Perret, X., Staehelin, C., Deakin, W.J., Raasch, C., Liesegang, H. and Broughton, W.J. (2004) An Evolutionary Hot Spot: the pNGR234b Replicon of Rhizobium sp. Strain NGR234. *Journal of Bacteriology*, **186**, 535-542.

Sullivan, J.T., Patrick, H.N., Lowther, W.L., Scott, D.B. and Ronson, C.W. (1995) Nodulating strains of Rhizobium loti arise through chromosomal symbiotic gene transfer in the environment. *Proceedings of the National Academy of Sciences*, **92**, 8985-8989.

Sullivan, J.T. and Ronson, C.W. (1998) Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. *Proceedings of the National Academy of Sciences*, **95**, 5145-5149.

Sullivan, J.T., Trzebiatowski, J.R., Cruickshank, R.W., Gouzy, J., Brown, S.D., Elliot, R.M., Fleetwood, D.J., McCallum, N.G., Rossbach, U., Stuart, G.S., Weaver, J.E., Webby, R.J., de Bruijn, F.J. and Ronson, C.W. (2002) Comparative Sequence Analysis of the Symbiosis Island of Mesorhizobium loti Strain R7A. *Journal of Bacteriology*, **184**, 3086-3095.

Suyama, M., Torrents, D. and Bork, P. (2006) PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research*, **34**, W609-W612.

Suyama, M., Torrents, D. and Bork, P. (2006) PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research*, **34**, W609-W612.

Tena, G., Boudsocq, M. and Sheen, J. (2011) Protein kinase signaling networks in plant innate immunity. *Current Opinion in Plant Biology*, **14**, 519-529.

Torres, M.A., Jones, J.D.G. and Dangl, J.L. (2006) Reactive Oxygen Species Signaling in Response to Pathogens. *Plant Physiology*, **141**, 373-378.

Triplett, E.W. and Sadowsky, M.J. (1992) Genetics of Competition for Nodulation of Legumes. *Annual Review of Microbiology*, **46**, 399-422.

Uchiumi, T., Ohwada, T., Itakura, M., Mitsui, H., Nukui, N., Dawadi, P., Kaneko, T., Tabata, S., Yokoyama, T., Tejima, K., Saeki, K., Omori, H., Hayashi, M., Maekawa, T., Sriprang, R., Murooka, Y., Tajima, S., Simomura, K., Nomura, M., Suzuki, A., Shimoda, Y., Sioya, K., Abe, M. and Minamisawa, K. (2004) Expression Islands Clustered on the Symbiosis Island of the Mesorhizobium loti Genome. *Journal of Bacteriology*, **186**, 2439-2448.

Underwood, W., Zhang, S. and He, S.Y. (2007) The Pseudomonas syringae type III effector tyrosine phosphatase HopAO1 suppresses innate immunity in Arabidopsis thaliana. *The Plant Journal*, **52**, 658-672.

van Berkum, P., Terefework, Z., Paulin, L., Suomalainen, S., LindstrÃ¶m, K. and Eardly, B.D. (2003) Discordant Phylogenies within the rrn Loci of Rhizobia. *Journal of Bacteriology*, **185**, 2988-2998.

van Rhijn, P. and Vanderleyden, J. (1995) The *Rhizobium*-Plant Symbiosis. *Microbiological Reviews*, **59**, 124-142.

Venkatesan, M.M., Goldberg, M.B., Rose, D.J., Grotbeck, E.J., Burland, V. and Blattner, F.R. (2001) Complete DNA Sequence and Analysis of the Large Virulence Plasmid of Shigella flexneri. *Infection and Immunity*, **69**, 3271-3285.

Vernikos, G.S. and Parkhill, J. (2008) Resolving the structural features of genomic islands: A machine learning approach. *Genome Research*, **18**, 331-342.

Viprey, V., Del Greco, A., Golinowski, W., Broughton, W.J. and Perret, X. (1998) Symbiotic implications of type III protein secretion machinery in Rhizobium. *Molecular Microbiology*, **28**, 1381-1389.

Viprey, V., Del Greco, A., Golinowski, W., Broughton, W.J. and Perret, X. (1998) Symbiotic implications of type III protein secretion machinery in Rhizobium. *Molecular Microbiology*, **28**, 1381-1389.

Wassem, R., Kobayashi, H., Kambara, K., Le Quéré, A., Walker, G.C., Broughton, W.J. and Deakin, W.J. (2008) TtsI regulates symbiotic genes in Rhizobium species NGR234 by binding to tts boxes. *Molecular Microbiology*, **68**, 736-748.

Wei, C.-F., Kvitko, B.H., Shimizu, R., Crabill, E., Alfano, J.R., Lin, N.-C., Martin, G.B., Huang, H.-C. and Collmer, A. (2007) A Pseudomonas syringae pv. tomato DC3000 mutant lacking the type III effector HopQ1-1 is able to cause disease in the model plant Nicotiana benthamiana. *The Plant Journal*, **51**, 32-46.

Weidner, S., Becker, A., Bonilla, I., Jaenicke, S., Lloret, J., Margaret, I., Puhler, A., Ruiz-Sainz, J.E., Schneiker-Bekel, S., Szczepanowski, R., Vinardell, J.M., Zehner, S. and Gottfert, M. (2012) Genome Sequence of the Soybean Symbiont Sinorhizobium fredii HH103. *Journal of Bacteriology*, **194**, 1617-1618.

Wenzel, M., Friedrich, L., Gottfert, M. and Zehner, S. (2009) The Type III-Secreted Protein NopE1 Affects Symbiosis and Exhibits a Calcium-Dependent Autocleavage Activity. *Molecular Plant-Microbe Interactions*, **23**, 124-129.

Willems, A. (2006) The taxonomy of rhizobia: an overview. *Plant and Soil*, **287**, 3-14.

Xiang, T., Zong, N., Zou, Y., Wu, Y., Zhang, J., Xing, W., Li, Y., Tang, X., Zhu, L., Chai, J. and Zhou, J.-M. (2008) Pseudomonas syringae Effector AvrPto Blocks Innate Immunity by Targeting Receptor Kinases. *Current biology : CB*, **18**, 74-80.

Xiao, Y., Lu, Y., Heu, S. and Hutcheson, S.W. (1992) Organization and environmental regulation of the Pseudomonas syringae pv. syringae 61 hrp cluster. *Journal of Bacteriology*, **174**, 1734-1741.

Xiao, Y., Heu, S., Yi, J., Lu, Y. and Hutcheson, S.W. (1994) Identification of a putative alternate sigma factor and characterization of a multicomponent regulatory cascade controlling the expression of Pseudomonas syringae pv. syringae Pss61 hrp and hrmA genes. *Journal of Bacteriology*, **176**, 1025-1036.

Yang, Z. and Nielsen, R. (2000) Estimating Synonymous and Nonsynonymous Substitution Rates Under Realistic Evolutionary Models. *Molecular Biology and Evolution*, **17**, 32-43.

Yang, F.-J., Cheng, L.-L., Zhang, L., Dai, W.-J., Liu, Z., Yao, N., Xie, Z.-P. and Staehelin, C. (2009) Y4lO of Rhizobium sp. Strain NGR234 Is a Symbiotic Determinant Required for Symbiosome Differentiation. *Journal of Bacteriology*, **191**, 735-746.

Yang, S., Tang, F., Gao, M., Krishnan, H.B. and Zhu, H. (2010) R gene-controlled host specificity in the legumeâ€“rhizobia symbiosis. *Proceedings of the National Academy of Sciences*, **107**, 18735-18740.

Yu, G.L., Katagiri, F. and Ausubel, F.M. (1993) Arabidopsis mutations at the RPS2 locus result in loss of resistance to Pseudomonas syringae strains expressing the avirulence gene avrRpt2. *Molecular Plant-Microbe Interactions*, **6**, 434-443.

Yuan, J. and He, S.Y. (1996) The Pseudomonas syringae Hrp regulation and secretion system controls the production and secretion of multiple extracellular proteins. *Journal of Bacteriology*, **178**, 6399-6402.

Zamioudis, C. and Pieterse, C.M.J. (2011) Modulation of Host Immunity by Beneficial Microbes. *Molecular Plant-Microbe Interactions*, **25**, 139-150.

Zerbino, D.R. and Birney, E. (2008) Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research*, **18**, 821-829.

Zhang, Z., Li, J., Zhao, X.Q., Wang, J., Wong, G.K. and Yu, J. (2006) KaKs_Calculator: calculating Ka and Ks through model selection and model averaging. *Genomics, Proteomics, & Bioinformatics*, **4**, 259-263.

Zhang, L., Chen, X.-J., Lu, H.-B., Xie, Z.-P. and Staehelin, C. (2011) Functional Analysis of the Type 3 Effector Nodulation Outer Protein L (NopL) from Rhizobium sp. NGR234. *Journal of Biological Chemistry*, **286**, 32178-32187.

Zhao, Y., Sundin, G.W. and Wang, D. (2009) Construction and analysis of pathogenicity island deletion mutants of Erwinia amylovora. *Canadian Journal of Microbiology*, **55**, 457-464.

Zhou, J.-M. and Chai, J. (2008) Plant pathogenic bacterial type III effectors subdue host responses. *Current Opinion in Microbiology*, **11**, 179-185.

Zhou, H., Morgan, R.L., Guttman, D.S. and Ma, W. (2009) Allelic Variants of the Pseudomonas syringae Type III Effector HopZ1 Are Differentially Recognized by Plant Resistance Systems. *Molecular Plant-Microbe Interactions*, **22**, 176-189.

Zuber, S., Carruthers, F., Keel, C., Mattart, A., Blumer, C., Pessi, G., Gigot-Bonnefoy, C.c., Schnider-Keel, U., Heeb, S., Reimmann, C. and Haas, D. (2003) GacS Sensor Domains Pertinent to the Regulation of Exoproduct Formation and to the Biocontrol Potential of Pseudomonas fluorescens CHA0. *Molecular Plant-Microbe Interactions*, **16**, 634-644.

Zwiesler-Vollick, J., Plovanich-Jones, A.E., Nomura, K., Bandyopadhyay, S., Joardar, V., Kunkel, B.N. and He, S.Y. (2002) Identification of novel hrp-regulated genes through functional genomic analysis of the Pseudomonas syringae pv. tomato DC3000 genome. *Molecular Microbiology*, **45**, 1207-1218.

APPENDICES

AutoSPOTs: Automated Image Analysis for Enumerating Callose Deposition

Jason S. Cumbie, Rebecca C. Pankow, William J. Thomas, and Jeff H. Chang

In *Genome-Enabled Analysis of Plant-Pathogen Interactions*
pp. 233-242

INTRODUCTION

Computational methods are essential to any genomicist's toolkit. With the continual advances in sequencing technology, there are demands for computational approaches that can keep pace with the different data structures. It is with these in mind that we have developed software programs to further enable integration of genomics with plant-pathogen research. In this chapter, we describe AutoSPOTs, one of the programs that we developed to facilitate high-throughput characterization of bacteria-plant interactions.

The type III secretion system (T3SS) is used by many Gram-negative bacteria to establish interactions with their hosts (Grant *et al*., 2006). The T3SS is a conduit that deploys bacterial encoded type III effector proteins directly into host cells where they function to manipulate the host for the benefit of the infecting bacterium. In the case of plant pathogenic bacteria, type III effectors are necessary to engage and dampen one layer of plant defense called PAMP-triggered immunity (PTI; Jones and Dangl, 2006). A number of events have been associated with PTI, including the deposition of callose in cell walls (Zipfel, 2009). Callose, a β-1,3 linked glucan, along with cellulose, pectin, lignin, and hydroxyproline-rich proteins, are deposited as an agglomeration believed to function as an apposition to infecting bacteria located in the apoplastic space and to other penetrating-type microbes (Bestwick *et al.*, 1995; Bestwick *et al.*, 1998).

Pseudomonas syringae is an excellent model pathogen of plants. The genome sequences for several strains of *P. syringae* have been completed and mined for candidate type III effector genes (Buell *et al.*, 2003; Feil *et al.*, 2005; Joardar *et al.*, 2005; Almeida *et al.*, 2009; Reinhardt *et al.*, 2009; Studholme *et al.*, 2009). Functional approaches that relied on the availability of the genome sequence have also been used (Chang et al., 2005). One strain in particular, *P. syringae* pv *tomato* race DC3000 (*Pto*DC3000), is intensively studied because of its ability to infect the model host plant, *Arabidopsis thaliana*. *Pto*DC3000 has approximately 30 type III effector genes (Schechter et al., 2006). The challenge now is to understand the functions of all type III effector

proteins and how a system of deployed type III effectors is coordinated in the host cell to dampen PTI for the benefit of the infecting bacterium.

AUTOSPOTS – FOR AUTOMATED BATCH ENUMERATION OF CALLOSE

Enumerating the deposition of callose is an often-used assay for quantifying PTI and perturbations to PTI. The wet-lab manipulations for this assay are relatively straightforward. The robustness of the assay, however, is affected by the variable host response to pathogen challenge and the obvious solution is to simply increase the number of samples. But, this simple solution is often outweighed by the onerous nature of the callose assay and its analyses.

We have therefore developed AutoSPOTs to mitigate the labor-intensive steps associated with image analyses and their potential associated biases. With user-defined criteria based on size and color, AutoSPOTs automates and batch enumerates aniline-stained callose deposits from JPEG images. AutoSPOTs will also automatically execute a series of standard statistical analyses. We have used AutoSPOTs to analyze thousands of images on a laptop computer. AutoSPOTs is an open-source Graphical User Interface (GUI) written in Perl and C. The software program and user's manual can be downloaded from our website at: http://changlab.cgrb.oregonstate.edu/.

Requirements for AutoSPOTs

Methods for sample preparation have been described (Kim and Mackey, 2008). Yet, some simple steps taken during sample preparation and microscopy can greatly improve the quality of the images for more accurate identification and enumeration of callose deposits. It is important to clear leaves as completely as possible subsequent to sample collection because autofluorescence of the chlorophyll will lead to background fluorescence. Insufficient staining can result in weakly fluorescent callose deposits. We have found that the simple act of staining leaves in aniline blue overnight improves the clarity of callose fluorescence. Proper mounting of leaves is another crucial step in sample preparation; wrinkling of leaves or bubbles in the mounting medium can result in

multiple focal planes in a single field of view, making resolution of the entire field difficult. Finally, it is important to use an appropriate exposure time for capturing high-quality images (this may require some trial and error). While the customizable color filter settings make AutoSPOTs functional over a range of exposures, extremes in exposures pose potential problems. Exposure settings that are too low will result in faint or dim callose deposits whereas exposure settings that are too high will wash out fluorescent spots. Both result in a reduction in the accuracy of AutoSPOTs.

We typically take ten JPEG images per leaf and sample fifteen leaves per treatment. A minimum of two treatments is required. For fully automated batch analysis, AutoSPOTs requires the user to properly name and store JPEG images in a recognizable manner. The two recognized formats are as single numbers (e.g., sample1.jpg, sample2.jpg), or as number-number (e.g., treatment1-1, treatment1-2). Additionally, there must be the same number of JPEG images per sample (leaf) per treatment group. JPEG images should be saved in directories labeled according to treatment groups. If these conditions are not met then some of the automated functions of AutoSPOTs cannot be used.

Defining filters

AutoSPOTs requires the user to define a size filter and one of two types of color filters. In a subsequent section of this chapter, we show the effects that different color filters have on results. AutoSPOTs will apply the filters on a pixel-by-pixel basis to identify callose deposits for each JPEG image to be analyzed. It is therefore important for the user to capture high quality JPEG images and to establish the proper filter settings for the most uniform, sensitive and accurate identification of aniline-blue stained callose deposits across an experiment.

For the size filter, we recommend starting with minimum and maximum sizes of 20 and 100, respectively, and to refine as needed (see discussion on previewing below). We have included two types of color filters: the RGB and ratio filters. To simplify selection, we have included a ‘color selection assistance’ feature. By selecting pixels of

callose deposits from several representative images, the color selection assistance feature will provide the user with the values for each of the criteria required of the RGB or ratio filters. Other criteria include 'Trip' and 'Drop' thresholds. The former is used by AutoSPOTs to determine which pixels will be considered as part of a stained callose deposit and 'trips' AutoSPOTs into expanding a callose deposit. The latter is used by AutoSPOTs to exclude pixels from a stained callose deposit and forces AutoSPOTs to 'drop' the pixel from expanding the callose deposit. The user can then determine the average values from multiple pixels of multiple images and set the color filters accordingly.

In most cases, AutoSPOTs performs better with grayscale JPEG images; this may depend on the camera and staining/de-staining of leaves. We have added a feature that enables all images to be automatically converted to grayscale. When defining the color filter, note that the red, green, and blue channels will have the same value so the ratio filter cannot be used. In contrast, the RGB filter must be used and simply becomes an RGB intensity filter.

AutoSPOTs allows the user to preview the sensitivity and accuracy of the filters. A screenshot of a preview and the GUI is presented (Fig. A1.1). The image will be displayed and each identified callose deposit will be demarked. The total number of callose deposits identified will also be displayed. It is strongly recommended that the user carefully examine several images and adjust the filter settings to find the desired level of sensitivity and accuracy. It is important to preview images with few and many callose deposits (see Fig. A1.3). We caution the user to pay close attention to identification of leaf features such as veins or trichomes as callose deposits as well as incomplete demarcation or over-extension of callose deposits. Incorrect identification of leaf features as callose deposits suggests the filters are too sensitive, whereas inaccurate demarcation of callose suggests the Trip and Drop distances are not correctly set.

It cannot be stressed enough that the successful use of AutoSPOTS will depend on consistent, high-quality images, control treatments to assess the accuracy of filters, application of filters uniformly on all samples of all treatments being compared, and a

sufficient number of JPEG images and samples to obtain good statistical power for analysis. Not all callose deposits will be identified, especially those in different focal planes, but as along as all leaves were prepared in a similar manner and JPEG images were photographed under similar settings, there will not be any biases in the results. We have provided a detailed step-by-step Users Manual available by download from our website.

Image analysis

Once the user has identified a satisfactory filter setting, AutoSPOTs can automatically batch process all images. Analysis begins by examining each pixel of each image individually to identify those that pass the 'Trip' threshold for a color filter. Once the pixels that pass the 'Trip' threshold are located, all adjoining pixels are analyzed using a 'Drop' threshold, which is usually a more relaxed threshold allowing for spot fading near the edges. Pixels are then continually counted outward until no more adjoining pixels can be found that match the 'Drop' threshold criteria. The number of pixels in a given 'spot' is tallied, and then analyzed using the size threshold. Those that are within the minimum and maximum values set by the user are counted as a single callose deposit.

AutoSPOTs calculates the average number of callose deposits by averaging per JPEG image per leaf per treatment. AutoSPOTS has built-in statistical analysis tools and will generate a statistical report for all treatments against the user-defined control treatment. AutoSPOTs will also plot the data for visual representation. At each step of analysis all the data is saved to text files and directories specified by the user. Copies of every image analyzed with their demarked callose deposits are also stored so the user can inspect the sensitivity and accuracy of the filters.

DEMONSTRATION OF AUTOSPOTS

We used one size filter setting and six different color filter settings in AutoSPOTs to demonstrate their effects on sensitivity and accuracy in enumerating callose deposits

from JPEG images (Fig. A1.2). Four of the tested color filter settings used RGB (intensity) values to identify callose deposits from JPEG images that were converted to grayscale. In these cases, the color filter setting was set from least sensitive to overly sensitive by using different values – we noted the drop and trip values had the largest effect on sensitivity. Two of the color filter settings used a ratio or RGB filter to analyze the original color JPEG images.

The treatments we tested were Arabidopsis infected with *Pto*DC3000, a T3SS-deficient mutant of *Pto*DC3000 (*hrcC*), a soil bacterium with an integrated T3SS-encoding region (EtHAn), and EtHAn carrying the type III effector gene, *hopM1*. *Pto*DC3000 deploys 30 type III effector proteins into Arabidopsis and sufficiently dampens PTI to cause disease. Its ability to dampen the deposition of callose has been repeatedly demonstrated (Hauck *et al*., 2003; DebRoy *et al*., 2004; Nomura *et al*., 2006; Ham *et al.*, 2007). In contrast, since the *hrcC* mutant is incapable of delivering type III effectors, it cannot dampen the deposition of callose or PTI, nor cause disease on Arabidopsis (Niepold *et al.*, 1985; Lindgren *et al.*, 1986; Roine *et al.*, 1997; Hauck *et al*., 2003; Thilmony *et al*., 2006). EtHAn was engineered from *P. fluorescens* Pf0-1 and is devoid of any endogenous type III effectors (Thomas *et al.*, 2009). EtHAn therefore elicits PTI. The type III effector, HopM1, is sufficient to dampen the deposition of callose (DebRoy *et al*., 2004; Nomura *et al*., 2006; Thomas *et al.*, 2009). A total of fifteen leaves were challenged per treatment, and ten images were randomly taken from each leaf. AutoSPOTs took less than 45 minutes to automatically analyze the 600 JPEG images.

In general, the trends were similar for each of the six filter settings (Fig. A1.2). However, when the automatically generated statistics were analyzed, it is clear that the filter settings do indeed affect interpretation of data. Based on previous findings, we expected significant differences between *Pto*DC3000 versus the *hrcC* mutant and EtHAn + *hopM1* versus EtHAn treatments. The color filter settings 1-3 resulted in no differences in the conclusions – both comparisons within each of the three settings were statistically significant. However, the color filter setting 1 was clearly the poorest of the three in terms of sensitivity. In contrast, increasing the sensitivity of grayscale analysis (setting 4)

or use of color JPEG images (setting 5 and 6) resulted in less desirable results. Thus, increased sensitivity to identify the highest number of callose deposits is not necessarily the most recommended approach.

We visually examined the analyzed JPEG images to understand the results of the different color filter settings (Fig. A1.3). In general, most of the color settings performed fairly well in analyzing areas with few callose deposits. Color filter settings 2, 3 and 6 were the more accurate. In contrast, the different color filter settings resulted in dramatic differences in the analysis of areas in which callose deposits were abundant. Settings 2 and 3 performed fairly well. However, very few callose deposits were identified in JPEG images with dense staining spots when AutoSPOTs used color filter settings 4 - 6. This was a consequence of AutoSPOTs failing to drop pixels and categorizing several callose deposits as one larger spot. These large spots would exceed the maximum of 100 as defined by the size filter and not be counted. Changing the size filter could potentially alleviate this problem to a certain extent. We have analyzed JPEG images provided by another research group and results from analysis of the color images were superior to grayscale images. This could be a consequence of differences in staining/de-staining of leaves or in the microscope camera. It is recommended to try different combinations of filters.

The differences in performance when analyzing JPEG images with sparse and dense callose deposits can lead to very misleading results. For example, we could not detect a significant difference between treatments with *Pto*DC3000 and its *hrcC* mutant under color filter setting numbers 4 and 5. This is because AutoSPOTs was sufficiently accurate in identifying the sparse callose deposits resulting from infection with *Pto*DC3000 but was inadequate in identifying densely distributed callose deposits resulting from infection with the *hrcC* mutant.

CONCLUSION

We developed AutoSPOTs a simple, user-friendly, and open-source software program to facilitate the high-throughput analysis of JPEG images. AutoSPOTs mitigates

labor-intensive data analysis by automating and batch analyzing large sets of JPEG images for callose deposits and comparing results between treatments. AutoSPOTs therefore provides the opportunity to examine larger numbers of type III effectors or host genetic backgrounds for their effects on PTI.

We purposefully developed AutoSPOTs to be a simple program. As a consequence, the filtering scheme that AutoSPOTs uses relies on the user to identify the most suitable combination of filters through careful visual examination of their JPEG images. It is therefore expected that the user will design a properly controlled experiment and capture high-quality and uniform JPEG images for analysis.

AutoSPOTs was developed for identification and enumeration of aniline-stained callose deposits but it has potential uses in other applications in studying plant-pathogen interactions, such as enumerating GFP-expressing bacteria.

ACKNOWLEDGEMENTS

We thank Caitlin A. Thireault, Allison Smith, and Philip Hillebrand for their assistance. We gratefully acknowledge Jim Carrington for use of his light microscope. This research was supported in part by start-up funds from Oregon State University to JHC and by the National Research Initiative Competitive Grant no. 2008-35600-18783 from the USDA National Institute of Food and Agriculture, Microbial Functional Genomics Program. JSC was supported by a Computational and Genome Biology Initiative Fellowship from Oregon State University.

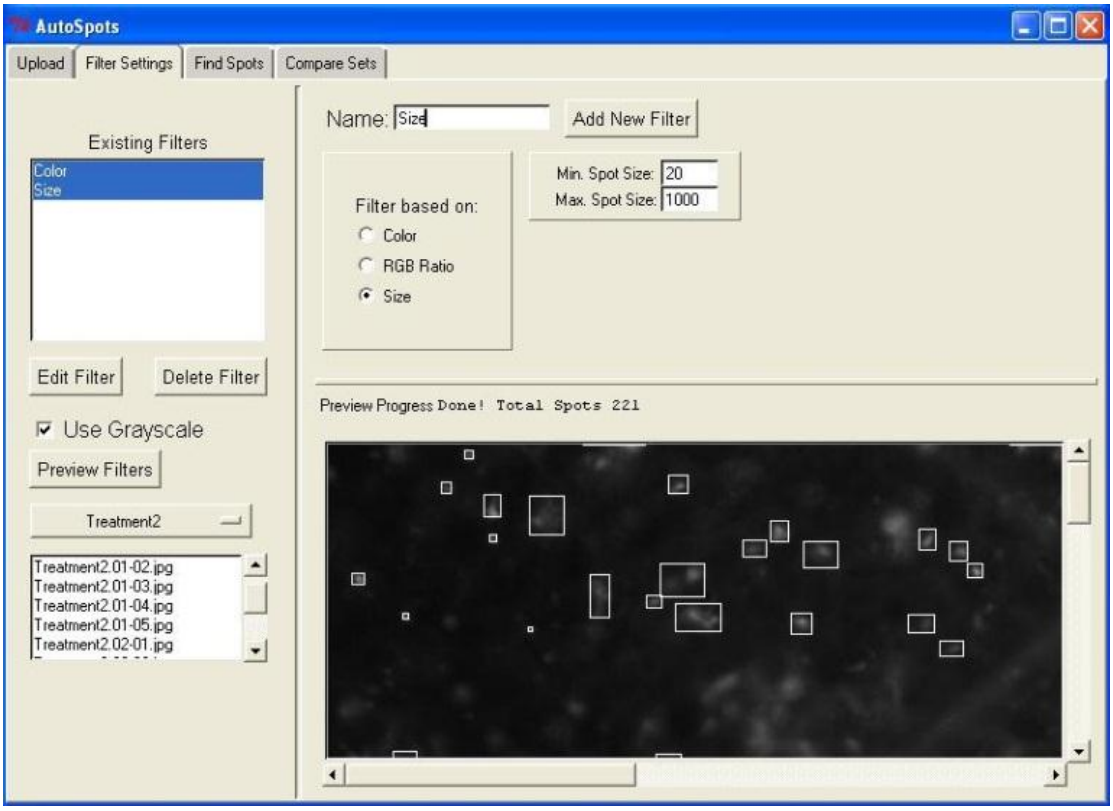


Figure A1.1. Screenshot of the Graphical User Interface of AutoSPOTs.
The AutoSPOTs GUI divides its various functions into four tabs. This screenshot of the Filter Settings tab illustrates the Preview Filters functions. Filters are defined and added in the top right section of the tab (settings for the Size filter are shown here). The desired filters are then selected from the Existing Filters menu (note that both color and size filters must be selected). The Use Grayscale option has been selected. Once a set of images has been loaded and an image selected in the lower left portion of the screen, the selected image will be displayed in the lower right display window. Callose deposits identified by the Preview Filters function will be indicated with a box and the total number of deposits identified will be displayed above the JPEG image.

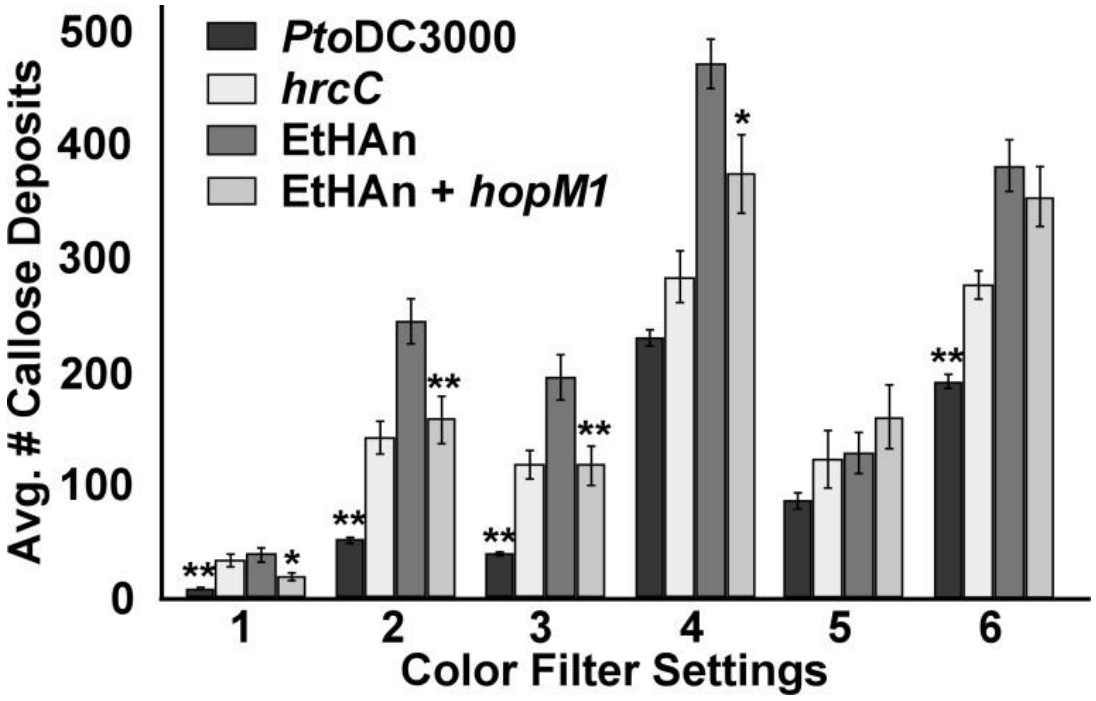


Figure A1.2. Enumeration of callose deposits by AutoSPOTs using different color filter settings.

We infected leaves of Arabidopsis with four different strains of bacteria. We used AutoSPOTs to identify and quantify callose deposits with six different color filter settings. For filters 1 – 4, JPEG images were converted to grayscale. The RGB, Trip, and Drop values respectively, were 130, 40, and 100 for color filter 1; 100, 80, and 100 for filter 2; 90, 50, and 80 for filter 3; and 90, 100, and 100 for color filter 4. For color filters 5 and 6, JPEG images were analyzed as color images using the color ratio and RGB filters, respectively. Fifteen leaves were infected per treatment and ten images were taken per leaf. Standard errors are shown. For each color filter setting, we compared results of *Pto*DC300 versus the *hrcC* mutant and EtHAn + *hopM1* versus EtHAn. Significant differences are denoted; *p-value ≤ 0.05; **p-value ≤ 0.01.

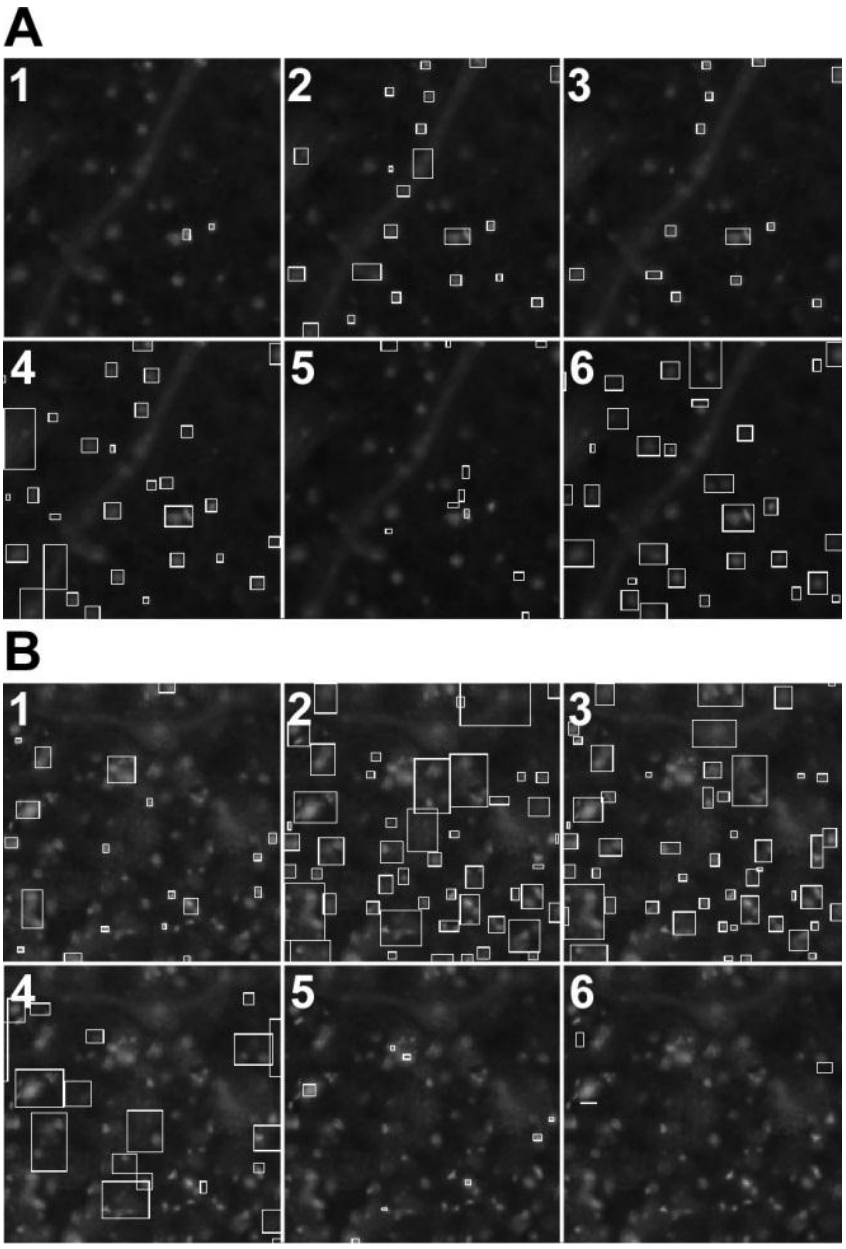


Figure A1.3. Effects of different color filter settings on the accuracy of AutoSPOTs.
Each set of panels represents the same section of the same two JPEG images analyzed using the six different color filter settings (1-6) described previously. One JPEG image had few callose deposits (A) while the other image was dense with callose deposits (B). Callose deposits identified by the program are indicated with a box. The analyses for color filters 5 and 6 used color images, which we have converted to grayscale for publication purposes.

REFERENCES CITED:

Almeida, N.F., Yan, S., Lindeberg, M., Studholme, D.J., Schneider, D.J., Condon, B., Liu, H., Viana, C.J., Warren, A., Evans, C., Kemen, E., Maclean, D., Angot, A., Martin, G.B., Jones, J.D., Collmer, A., Setubal, J.C., and Vinatzer, B.A. 2009. A draft genome sequence of *Pseudomonas syringae* pv. *tomato* T1 reveals a type III effector repertoire significantly divergent from that of *Pseudomonas syringae* pv. *tomato* DC3000. Mol. Plant-Microbe Interact. 22:52-62.

Bestwick, C.S., Bennett, M.H., and Mansfield, J.W. 1995. *Hrp* mutant of *Pseudomonas syringae* pv *phaseolicola* induces cell wall alterations but not membrane damage leading to the hypersensitive reaction in lettuce. Plant Physiol. 108:503-516.

Bestwick, C.S., Brown, I.R., and Mansfield, J.W. 1998. Localized changes in peroxidase activity accompany hydrogen peroxide generation during the development of a nonhost hypersensitive reaction in lettuce. Plant Physiol. 118:1067-1078.

Buell, C.R., Joardar, V., Lindeberg, M., Selengut, J., Paulsen, I.T., Gwinn, M.L., Dodson, R.J., Deboy, R.T., Durkin, A.S., Kolonay, J.F., Madupu, R., Daugherty, S., Brinkac, L., Beanan, M.J., Haft, D.H., Nelson, W.C., Davidsen, T., Zafar, N., Zhou, L., Liu, J., Yuan, Q., Khouri, H., Fedorova, N., Tran, B., Russell, D., Berry, K., Utterback, T., Van Aken, S.E., Feldblyum, T.V., D'Ascenzo, M., Deng, W.L., Ramos, A.R., Alfano, J.R., Cartinhour, S., Chatterjee, A.K., Delaney, T.P., Lazarowitz, S.G., Martin, G.B., Schneider, D.J., Tang, X., Bender, C.L., White, O., Fraser, C.M., and Collmer, A. 2003. The complete genome sequence of the Arabidopsis and tomato pathogen *Pseudomonas syringae* pv. *tomato* DC3000. Proc. Natl. Acad. Sci. U S A 100:10181-10186.

Chang, J.H., Urbach, J.M., Law, T.F., Arnold, L.W., Hu, A., Gombar, S., Grant, S.R., Ausubel, F.M., and Dangl, J.L. 2005. A high-throughput, near-saturating screen for type III effector genes from *Pseudomonas syringae*. Proc. Natl. Acad. Sci. U S A 102:2549-2554.

DebRoy, S., Thilmony, R., Kwack, Y.B., Nomura, K., and He, S.Y. 2004. A family of conserved bacterial effectors inhibits salicylic acid-mediated basal immunity and promotes disease necrosis in plants. Proc. Natl. Acad. Sci. U S A 101:9927-9932.

Feil, H., Feil, W.S., Chain, P., Larimer, F., DiBartolo, G., Copeland, A., Lykidis, A., Trong, S., Nolan, M., Goltsman, E., Thiel, J., Malfatti, S., Loper, J.E., Lapidus, A., Detter, J.C., Land, M., Richardson, P.M., Kyrpides, N.C., Ivanova, N., and Lindow, S.E. 2005. Comparison of the complete genome sequences of *Pseudomonas syringae* pv. *syringae* B728a and pv. *tomato* DC3000. Proc. Natl. Acad. Sci. U S A 102:11064-11069.

Grant, S.R., Fisher, E.J., Chang, J.H., Mole, B.M., and Dangl, J.L. 2006. Subterfuge and manipulation: Type III effector proteins of phytopathogenic bacteria. Ann. Rev. Microbiol. 60:425-449.

Ham, J.H., Kim, M.G., Lee, S.Y., and Mackey, D. 2007. Layered basal defenses underlie non-host resistance of Arabidopsis to *Pseudomonas syringae* pv. *phaseolicola*. Plant J. 51:604-616.

Hauck, P., Thilmony, R., and He, S.Y. 2003. A *Pseudomonas syringae* type III effector suppresses cell wall-based extracellular defense in susceptible Arabidopsis plants. Proc. Natl. Acad. Sci. U S A 100:8577-8582.

Joardar, V., Lindeberg, M., Jackson, R.W., Selengut, J., Dodson, R., Brinkac, L.M., Daugherty, S.C., Deboy, R., Durkin, A.S., Giglio, M.G., Madupu, R., Nelson, W.C., Rosovitz, M.J., Sullivan, S., Crabtree, J., Creasy, T., Davidsen, T., Haft, D.H., Zafar, N., Zhou, L., Halpin, R., Holley, T., Khouri, H., Feldblyum, T., White, O., Fraser, C.M., Chatterjee, A.K., Cartinhour, S., Schneider, D.J., Mansfield, J., Collmer, A., and Buell, C.R. 2005. Whole-genome sequence analysis of *Pseudomonas syringae* pv. *phaseolicola* 1448A reveals divergence among pathovars in genes involved in virulence and transposition. J. Bacteriol. 187:6488-6498.

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

Kim, M.G., and Mackey, D. 2008. Measuring cell-wall-based defenses and their effect on bacterial growth in Arabidopsis. Methods Mol. Biol. 415:443-452.

Lindgren, P.B., Peet, R.C., and Panopoulos, N.J. 1986. Gene cluster of *Pseudomonas syringae* pv. "*phaseolicola*" controls pathogenicity of bean plants and hypersensitivity of nonhost plants. J. Bacteriol. 168:512-522.

Niepold, F., Anderson, D., and Mills, D. 1985. Cloning determinants of pathogenesis from *Pseudomonas syringae* pathovar *syringae*. Proc. Natl. Acad. Sci. USA 82:406-410.

Nomura, K., Debroy, S., Lee, Y.H., Pumplin, N., Jones, J., and He, S.Y. 2006. A bacterial virulence protein suppresses host innate immunity to cause plant disease. Science 313:220-223.

Reinhardt, J.A., Baltrus, D.A., Nishimura, M.T., Jeck, W.R., Jones, C.D., and Dangl, J.L. 2009. De novo assembly using low-coverage short read sequence data from the rice pathogen *Pseudomonas syringae* pv. *oryzae*. Genome Res. 19:294-305.

Roine, E., Wei, W., Yuan, J., Nurmiaho-Lassila, E.L., Kalkkinen, N., Romantschuk, M., and He, S.Y. 1997. Hrp pilus: an hrp-dependent bacterial surface appendage produced by *Pseudomonas syringae* pv. *tomato* DC3000. Proc. Natl. Acad. Sci. U S A 94:3459-3464.

Schechter, L.M., Vencato, M., Jordan, K.L., Schneider, S.E., Schneider, D.J., and Collmer, A. 2006. Multiple approaches to a complete inventory of *Pseudomonas syringae* pv. *tomato* DC3000 type III secretion system effector proteins. Mol. Plant-Microbe Interact. 19:1180-1192.

Studholme, D.J., Ibanez, S.G., MacLean, D., Dangl, J.L., Chang, J.H., and Rathjen, J.P. 2009. A draft genome sequence and functional screen reveals the repertoire of type III secreted proteins of *Pseudomonas syringae* pathovar *tabaci* 11528. BMC Genomics 10:395.

Thilmony, R., Underwood, W., and He, S.Y. 2006. Genome-wide transcriptional analysis of the *Arabidopsis thaliana* interaction with the plant pathogen *Pseudomonas syringae* pv. *tomato* DC3000 and the human pathogen *Escherichia coli* O157:H7. Plant J. 46:34-53.

Thomas, W.J., Thireault, C.A., Kimbrel, J.A., and Chang, J.H. 2009. Recombineering and stable integration of the *Pseudomonas syringae* pv. *syringae* 61 *hrp/hrc* cluster into the genome of the soil bacterium *Pseudomonas fluorescens* Pf0-1. Plant J. 60:919-928.

Zipfel, C. 2009. Early molecular events in PAMP-triggered immunity. Curr. Opin. Plant Biol. 12:414-420.

Integration of a repertoire of type III effector genes from *Pseudomonas syringe* pv. *tomato* DC3000 into the genomes of *Pseudomonas syringae* pv. *phaseolicola* 1448a and EtHAn

William J. Thomas, Jeff H. Chang

Preliminary results for a project in progress

INTRODUCTION

Plant-associated bacteria face formidable barriers to establishing symbioses with their host organisms. Like other multicellular eukaryotic organisms, plants possess innate immunity that programs host cells to detect and respond to microbial attack (Nürnberger *et al.*, 2004; Postel and Kemmerling, 2009). Recognition of microbe-associated signals induces expression of defensive responses that provide an effective barrier to symbioses with most microbes. Many successful microsymbionts circumvent these barriers through the delivery of microbe-encoded effector proteins into host cells, where they manipulate the host's cellular machinery to dampen immune responses (Postel and Kemmerling, 2009; Jones and Dangl, 2006). Typically, this suppression of host immunity requires the activity of multiple effector proteins; for this reason, microsymbionts that use a secretion system for effector delivery typically encode a collection of compatible effector proteins (Galán and Wolf-Watz, 2006). The size and makeup of these effector protein collections varies depending on the lifestyle of the microsymbiont, which may range from mutualism to parasitism, and the biology of the host organism. Thus, although effector repertoires may overlap to some degree, a given microsymbiont will have a unique collection suited to establishing a specific host-microbe interaction (Grant *et al.*, 2006).

Many Gram-negative bacteria encode a set of effectors that are delivered via the type III secretion system (T3SS). Among plant-associated bacteria, T3SSs have been identified in the nitrogen-fixing mutualist rhizobia, as well as many environmentally and agriculturally important phytopathogens. Previous published work has primarily focused on the distribution of T3Es of the latter group, wherein they are widely distributed among many species of economically important agricultural pests. In this context, T3E repertoires have proven to be staggeringly diverse between species of phytopathogenic bacteria, with enormous variation even between strains of the same species (Baltrus *et al*., 2011). The content of these repertoires appears to be influenced by the pathogenic lifestyle of a given bacterium; ranging from collections of T3Es with as few as three effectors in necrotrophic *Erwinia amylovora*, to large collections of dozens of T3Es in

hemibiotrophic *Ralstonia solanacearum* and *Pseudomonas syringae* (Oh and Beer, 2005; Cunnac *et al.*, 2004; Chang *et al.*, 2005).

In phytopathogenic bacteria that encode T3E repertoires, the T3Es are collectively necessary for full virulence on host plants (Cunnac *et al.*, 2009; Lewis *et al.*, 2009). T3Es are hypothesized to suppress the innate immunity of host plant cells below the threshold required for resistance to infection (Jones and Dangl, 2006). In plants, innate immunity is mediated by the recognition of conserved microbe-associated molecular patterns (MAMPs) via host-encoded pattern recognition receptors (PRRs; Sanabria *et al.*, 2010; Segonzac and Zipfel, 2011). MAMP-derived signals are transduced by their cognate PRRs into a battery of immune responses collectively referred to as MAMP-triggered immunity, or MTI (Jones and Dangl, 2006; Schwessinger and Zipfel, 2008). MTI initiates with responses that include a MAP kinase cascade that culminates in changes to transcriptional regulation, Ca^{2+} influx, rapid oxidative burst, and nitric oxide signaling (Schwessinger and Zipfel, 2008; Pitzschke *et al.*, 2009; Tena *et al.,* 2011; Segonzac and Zipfel, 2011; Ma, 2011). Subsequent responses involve activation of regulatory phytohormones including salicylic acid, ethylene, and jasmonic acid, which regulate cellular and systemic responses to microbial invasion (Pitzschke *et al.*, 2009; Segonzac and Zipfel, 2011; Tena *et al.,* 2011; An and Mou, 2011; Robert-Seilaniantz *et al.*, 2011). These outputs of MTI are sufficient for resistance to most microsymbionts; however, bacteria with a compatible collection of T3Es are able to circumvent these barriers to infection (Grant *et al.*, 2006; Lewis *et al.*, 2009).

Although collections of T3Es serve as positive determinants of host range in their role as suppressors of MTI, individual T3Es may function as negative host-range determinants due to their potential to elicit another layer of plant immunity (Jones and Dangl, 2006; Zhou, 2008; Mansfield, 2009; Rafiqi *et al.*, 2009). Plants have specialized resistance (R) proteins that perceive specific T3Es, eliciting a battery of responses that resembles an accelerated, amplified MTI, which frequently culminates in programmed cell death, referred to in this context as hypersensitive response (HR). T3Es that are recognized by R proteins have historically been designated as avirulence (Avr) proteins

because of their limiting effect on host range. Therefore, in the context of T3E-mediated symbioses, a compatible host-microbe interaction requires a collection of T3Es that is sufficient to dampen MTI while avoiding induction of ETI.

In a paradigm in which T3E repertoires are the primary host-range determinants in plant-microbe interactions, two important hypotheses arise. If a collection of T3Es is sufficient to dampen MTI below the threshold of effective resistance, one can hypothesize that the transfer of an entire collection of T3E from a phytopathogen to a non-pathogenic bacterium will convert the non-pathogen into a host-associated microsymbiont specific to the host of the phytopathogen from which the T3Es were derived. Secondly, a collection of T3Es from a phytopathogen will have the potential to upset the host range of another T3SS-encoding phytopathogen and expand it to include the host plant that is compatible with the pathogen from which the T3Es were derived.

Here, I describe the initial steps to test these two hypotheses. To address the hypothesis that a complete set of T3Es is necessary and sufficient for pathogenesis, I transferred a near-complete T3E gene repertoire of the tomato and Arabidopsis pathogen *Pseudomonas syringae* pv. *tomato* DC3000 (*Pto*DC3000) into *Pseudomonas fluorescens* Pf0-1 that had been previously modified to encode an exogenous T3SS from another *P. syringae* strain (EtHAn, see Chapter 4; Thomas *et al.*, 2009). If a T3E repertoire was sufficient for pathogenesis, then the modified EtHAn would gain the ability to infect a compatible host of *Pto*DC3000. To test the second hypothesis, I also cloned the same collection of 28 *Pto*DC3000 T3E genes into *P. syringae* pv. *phaseolicola* 1448a (*Pph*1448a). *Pph*1448a is a pathogen of beans but is not adapted to Arabidopsis, and triggers PTI and a so-called weak ETI (Ham et al., 2007). Nonetheless, when co-inoculated with PtoDC3000, but not its T3SS-deficient mutant, Pph1448a could achieve high levels of growth in planta. This observation suggests the T3E collection of PtoDC3000 could sufficiently dampen both PTI and weak ETI caused by Pph1448.

In this appendix, I describe the cloning of a collection of 28 T3E genes from *Pto*DC3000 into EtHAn and *Pph*1448a, as well as preliminary investigation into the

sufficiency of this set of T3S in changing their host ranges. The experiments described here were knowingly high risk/high reward.

RESULTS

The *Pto*DC3000 T3E repertoire is not sufficient for EtHAn to be pathogenic on Arabidopsis

To test the hypothesis that a full repertoire of T3Es from a phytopathogen would suffice to enable a non-pathogenic strain to be pathogenic, I quantified the ability of EtHAn carrying the T3E genes of *Pto*DC3000, to grow *in planta*. *Pto*DC3000, in a T3SS- and T3E-dependent manner, can grow to high levels *in planta* within 3 days post-inoculation (dpi) and elicit disease, characterized by visible tissue collapse and necrosis, approximately 28 hours post-inoculation (hpi). I infiltrated three variants of EtHAn carrying the *Pto*DC3000 T3E repertoire ($EtHAn_{PtoDC3000}$), in which each clone had been integrated independently of the other variants, into Col-0 leaves at a concentration of 1.0×10^8 cfu/ml. After ~28 hpi, 100% of the leaves inoculated with wild-type *Pto*DC3000 exhibited tissue collapse characteristic of disease, while none of the variants of $EtHAn_{PtoDC3000}$ had responded, a phenotype indistinguishable from EtHAn without T3E genes (Figure A2.2). At 72 hpi, the observed phenotypes remained unchanged, with the T3E-equipped $EtHAn_{PtoDC3000}$ strains still exhibiting an identical phenotype to the negative control.

I also quantified the growth of $EtHAn_{PtoDC3000}$ *in planta*. I infiltrated Arabidopsis Col-0 leaves with $EtHAn_{PtoDC3000}$ at a lower dose and assessed growth over a seven-day period. No increased growth was evident in $EtHAn_{PtoDC3000}$-treated leaves compared to those treated with unmodified EtHAn (Figure A2.3).

The *Pto*DC3000 T3E repertoire is not sufficient for *Pph*1448a to be pathogenic on Arabidopsis

To test the hypothesis that the T3E repertoire of *Pto*DC3000 would modify the host range of the bean pathogen, *Pph*1448a, I carried out parallel experiments as described above. After infiltrating Arabidopsis Col-0 leaves with high-dose inocula of

*Pph*1448a equipped with the *Pto*DC3000 T3E repertoire (*Pph*1448a$_{PtoDC3000}$), control leaves treated with *Pto*DC3000 exhibited the tissue collapse characteristic of disease, while leaves infected with *Pph*1448a$_{PtoDC3000}$ showed no visible phenotype, mirroring the phenotype observed in leaves treated with wild type *Pph*1448a (Figure A2.4). When assessed at 72 hpi, both wild type *Pph*1448a- and *Pph*1448a$_{PtoDC3000}$-treated leaves remained identical in phenotype, with neither treatment eliciting disease symptoms.

Enumeration of *in planta* growth of *Pph*1448a$_{PtoDC3000}$ in Arabidopsis leaves infiltrated with low-dose inocula showed no evidence of increased growth of *Pph*1448a$_{PtoDC3000}$ compared to wild type *Pph*1448a (Figure A2.5).

SUMMARY

The current paradigm of T3E-mediated host-bacteria interactions predicts that a collection of T3Es encoded by a bacterium is sufficient to dampen MTI below the threshold of effective resistance to infection (Grant, 2006). Implicit in this paradigm is the assumption that an organism that dampens MTI can successfully establish a host-microbe relationship with the immune-suppressed plant cell. Based on this, we first tested the null hypothesis that collection of T3Es, which is sufficient for a phytopathogen to cause disease in its host, will be sufficient to enable a non-pathogenic bacterium to suppress MTI in the original phytopathogen's host and proliferate in its tissue. An alternative, and more likely, prediction is that pathogens require additional virulence factors to successfully pathogenize its host. Thus, we also tested the hypothesis that a T3E collection from one phytopathogen will confer its host specificity to another non-adapted pathogen. To test these hypotheses, we cloned a 28-T3E gene repertoire from the tomato and Arabidopsis pathogen *Pto*DC3000 into the modified soil bacterium EtHAn and the bean pathogen *Pph*1448a.

The project described in this appendix was inherently risky. However, it has the potential to make a significant contribution to our understanding of the role of T3Es in mediating host-microbe relationships. Unfortunately, the T3E genes of PtoDC3000 were insufficient to confer virulence to either EtHAn or Pph1448. However, the results

described herein are far from complete and significant work remains. At this point, it is unclear whether the transferred T3Es are expressed in the heterologous strains. Translocation of the heterologous T3Es has not yet been demonstrated; this could be inferred from HR assays dependent on delivery of exogenous Avr proteins from EtHAn$_{Pto\text{DC3000}}$ and *Pph*1448a$_{Pto\text{DC3000}}$. Furthermore, to address the possibility of disruption of one or more important genes by the Tn5-mediated integration of the *Pto*DC3000 T3E repertoire, additional alternate versions of *Pph*1448a$_{Pto\text{DC3000}}$ should be completed, in which each Tn5-mediated integration has occurred independently of other versions.

Given the high-risk nature of this large undertaking, it is perhaps understandable that I undertook the "high-reward" experiments that, if successful, would have offered immediate support for the hypotheses being investigated. While these very preliminary results are only the very beginnings of the answer to our research questions, they offer a hint of what might be learned about the role of collections of T3E effectors in plant-bacterial symbioses.

MATERIALS AND METHODS

Bacterial strains, plant lines and growth conditions

The bacterial strains used in this study are *Pto*DC3000, *Pph*1448a, and *P. fluorescens* Pf0-1 modified with a stably integrated *P. syringae* pv. *syringae* 61 T3SS locus (Silby *et al.*, 2009; see Chapter 4; Thomas, *et al.*, 2009), *Escherichia coli* DH5a and HB101lpir. Pseudomonads were grown at 28°C in King's B (KB) liquid media with shaking or on KB agar plates. Antibiotics were used at the final concentrations of: 25 mg/ml rifampicin, 30 mg/ml kanamycin (100 mg/ml for Pf0-1), 30 mg/ml chloramphenicol, and 25 mg/ml gentamycin (100 mg/ml for Pf0-1). Concentrations listed were for *P. syringae* and *E. coli* unless otherwise noted.

Arabidopsis thaliana accession Col-0 plants were grown in a controlled-environment growth chamber (9 hrs of day at 22°C, 15 hrs of night at 20°C) for 5~6 weeks.

Plasmid constructions

Plasmids used were pBBR1-MCS1 (Kovach *et al.*, 1994), pRK2013 (Figurski and Helinski, 1979), mini-*Tn5* (de Lorenzo *et al.*, 1990), pKD4 and pKD46 (Datsenko and Wanner, 2000), pBH474 (House *et al.*, 2004), pME3280a (Zuber *et al.*, 2003), pUX-BF13 (Bao *et al.*, 1991), and pLN18 (Jamir *et al.*, 2004). Sequences of oligonucleotides used are available upon request. All restriction enzymes were purchased from New England Biolabs (NEB; Ipswich, MA).

Clone 1:

We used recombinant and sticky-end PCR to fuse together two 50 bp-sized fragments encoding the 5' and 3' regions of the full-length genes *hopAB2, hopI1, avrE1, shcM-hopM1,* and *avrE*, respectively. The 50-bp fragments from the 5' end and 3' end of each gene were amplified in separate reactions using Pfu (a list of primers used is available upon request). The products were gel-purified. Upstream and downstream fragments were mixed in approximately equal ratios and each amplified using the top primer from the reaction that generated the 5' fragment, combined with the bottom primer from the reaction that produced the 3' fragment. Recombined products carried inserted unique restriction sites: *hopAB2* flanking regions were separated by and *Xba*I restriction site, *hopI*I by an *Nhe*I restriction site, *avre1* by an *Acc*65I restriction site, *shcM-hopM1* by an *Xho*I restriction site, and *shcE* by a *Spe*I restriction site. The products of this PCR were gel-purified. The 100-bp products of *hopAB2* and *hopI1* were mixed in approximately equal ratios and amplified using the top primer for the 5' end of *hopAB2* and the bottom primer from the 3' end of *hopI1*. The 100-bp products of *avrE1* and *shcM-hopM1* were mixed in approximately equal ratios and amplified using the top primer for the 5' end of *avrE1* and the bottom primer from the 3' end of *shcM-hopM1*.

We also used recombinant and sticky-end PCR to fuse together the full-length genes *hopAM1-1* and *avrPto1.* Both genes were amplified in separate reactions using Pfu, and the products were gel-purified. The gel-purified products were mixed in

approximately equal ratios and amplified using the top primer for *hopAM1-1* and the bottom primer from *avrPto1.*

All products were gel-purified, and the sticky-end products from the fused flanking regions of *hopAB2* and *hopI1* cloned into pBBR1-MCS1 cleaved with *Bam*HI and *Xba*I. This vector was then cleaved with *Acc*65I and *Xho*I and the fusion of full-length *hopAM1-1* and *avrPto1* with sticky ends was ligated in. We then cleaved this vector using *Xho*I and *Eco*RV and cloned in the sticky-end products from the fused flanking regions of *avrE1* and *shcM-hopM1-1*. Finally, we cleaved the vector again with *Eco*RI and *Hind*III and cloned the sticky-end product from the fused flanking regions of *shcE*.

Clone 2:

We used recombinant and sticky-end PCR to fuse together two 50 bp-sized fragments encoding the 5' and 3' regions of the operons *shcF2-hopF2-hopU1* and *shcO1-hopO1-1-hopT1*, and the full-length genes *hopH1, hopAO1,* and *hopH1,* respectively. The recombinant sticky-end PCR products were generated as described for clone 1 above. Recombined products carried inserted unique restriction sites: *hopH1* flanking regions were separated by an *Eco*RV restriction site, *hopAO1* by a *Bam*HI restriction site, *hopD1* by a *Sma*I restriction site, *shcF2-hopF2-hopU1* by a *Cla*I restriction site, and *shcO1-hopO1-1-hopT1* by an *Acc*65I restriction site. The products of this PCR were gel-purified. The 100-bp products of *shcF2-hopF2-hopU1* and *shcO1-hopO1-1-hopT1* were mixed in approximately equal ratios and amplified using the top primer for the 5' end of *shcF2-hopF2-hopU1* and the bottom primer from the 3' end of *shcO1-hopO1-1-hopT1*. The 100-bp products of *hopH1* and *hopAO1* were mixed in approximately equal ratios and amplified using the top primer for the 5' end of *hopH1*and the bottom primer from the 3' end of *hopAO1* .

We also used recombinant and sticky-end PCR to fuse together the full-length genes *hopG1* and *hopC1.* Both genes were amplified in separate reactions using Pfu, and the products were gel-purified. The gel-purified products were mixed in approximately

equal ratios and amplified using the top primer for *hopG1* and the bottom primer from *hopC1.*

All products were gel-purified, and the sticky-end products from the fused flanking regions of *shcF2-hopF2-hopU1* and *shcO1-hopO1-1-hopT1* were cloned into pBBR1-MCS1 cleaved with *Acc*65I and *Hind*III. This vector was then cleaved with *Bam*HI and *Xba*I and the sticky-end product from the fused flanking regions of *hopD1* was ligated in. We then cleaved this vector using *Hind*III and *Eco*RV and cloned in the fusion of full-length *hopG1* and *hopC1* with sticky ends. Finally, we cleaved the vector again with *Eco*RV and *Bam*HI and cloned the sticky-end product from the fused flanking regions of *hopH1* and *hopAO1.*

Clone 3:

We used recombinant and sticky-end PCR to fuse together two 50 bp-sized fragments encoding the 5' and 3' regions of the full-length genes *hopAA1-1,* and *hopR1,* respectively. The recombinant sticky-end PCR products were generated as described for clone 1 above. Recombined products carried inserted unique restriction sites: *hopR1* flanking regions were separated by a *Bam*HI restriction site and *hopAA1-1* by an *Xba*I restriction site. The products of this PCR were gel-purified. The 100-bp products of *hopAA1-1* and *hopR1* were mixed in approximately equal ratios and amplified using the top primer for the 5' end of *hopAA1-1* and the bottom primer from the 3' end of *hopR1*.

We also used recombinant and sticky-end PCR to fuse together the full-length genes *hopP1* and *hopY1*, and to fuse together the full-length genes *hopE1* and *hopAM1-2.* Genes were amplified in separate reactions using Pfu, and the products were gel-purified. The gel-purified products of *hopP1* and *hopY1* were mixed in approximately equal ratios and amplified using the top primer for *hopP1* and the bottom primer from *hopY1.* The gel-purified products of *hopE1* and *hopAM1-2* were mixed in approximately equal ratios and amplified using the top primer for *hopE1* and the bottom primer from *hopAM1-2.* Sticky-end PCR was used to generate sticky-end products for full-length *hopK1* and *shcN1-hopN1.*

All products were gel-purified, and the full-length fusion of *hopE1* and *hopAM1-2* were cloned into pBBR1-MCS1 cleaved with *Acc*65I and *Xho*I. This vector was then cleaved with *Xho*I and *Hind*III and the sticky-end product of full-length *shcN1-hopN1* was ligated in. We then cleaved this vector using *Hind*III and *Eco*RV and cloned in the sticky end product of full-length *hopK1*. Next, we cleaved the vector with *Eco*RV and *Bam*HI and cloned in the sticky-end product from the fused full-length *hopP1* and *hopY1.* Finally, we cleaved the vector again with *Bam*HI and *Xba*I and cloned in the sticky-end products from the fused flanking regions of *hopAA1-1* and *hopR1*.

Clone 4:

We used recombinant and sticky-end PCR to fuse together two 50 bp-sized fragments encoding the 5' and 3' regions of the operons *shcV1-hopV1* and *shcA1-hopA1*, and the full-length gene *hopQ1-1,* respectively. The recombinant sticky-end PCR products were generated as described for clone 1 above. Recombined products carried inserted unique restriction sites: *shcV1-hopV1* flanking regions were separated by an *Xba*I restriction site, *hopQ1-1* by a *Sma*I restriction site, and *shcA1-hopA1* by a *Bgl*II restriction site. The products of this PCR were gel-purified. The 100-bp products of *shcV1-hopV1* and *hopQ1-1* were mixed in approximately equal ratios and amplified using the top primer for the 5' end of *shcV1-hopV1* and the bottom primer from the 3' end of *hopQ1-1*.

We also used recombinant and sticky-end PCR to fuse together the full-length genes *hopAF1* and *hopX1.* Both genes were amplified in separate reactions using Pfu, and the products were gel-purified. The gel-purified products were mixed in approximately equal ratios and amplified using the top primer for *hopAF1* and the bottom primer from *hopX1.* Sticky-end PCR was used to generate sticky-end products for full-length *hopAA1-2.*

All products were gel-purified, and the sticky-end products from the fused flanking regions of *shcV1-hopV1* and *hopQ1-1* were cloned into pBBR1-MCS1 cleaved with *Bam*HI and *Xba*I. This vector was then cleaved with *Hind*III and *BamH*I and the

sticky-end product of full-length *hopAA1-2* was ligated in. We then cleaved this vector using *Acc*65I and *Xho*I and cloned in the fused flanking regions of *hopAF1* and *hopX1*. Finally, we cleaved the vector again with *Xho*I and *Hind*III and cloned the sticky-end product from the fused flanking regions of *shcA1-hopA1.*

Recombineering

Sequential recombineering steps for each clone were done in the order indicated in Figure A2.1. To recombineer a full-length gene, the pBBR1-MCS1 clone was digested at the unique restriction site separating the flanking regions of that gene. The linearized pBBR1-MCS1 clone was then transformed into electrocompetent cells made from arabinose-induced (10 mM) HB101λpir cells carrying pKD46 and the full-length gene encoded on pDONR207 (Invitrogen). Transformants were selected on chloramphenicol. Successful transfer of the full-length gene was confirmed by PCR.

We next cleaved the modified mini-*Tn5* with FRT-flanked Kan^r encoding flanking regions from pBBR1-MCS1 (see chapter 4; Thomas *et al.*, 2009) with *Acc*65I and transformed the gel-purified, linear fragment into electrocompetent cells made from arabinose-induced (10 mM) HB101λpir cells carrying pKD46 and pBBR1-MCS1 with the T3E-encoding region. Transformants were selected on kanamycin. Successful transfer of the T3SS-encoding region was confirmed by PCR.

Plasmid mobilization

Plasmids were mobilized into recipients as previously described (Chang *et al.*, 2005). Clones 1-4 were integrated into the target genomes sequentially, necessitating eviction of the Kan^r cassette subsequent to each Tn5-mediated integration that preceded the integration of another set of T3Es. Eviction of Kan^R was mediated by pBH474, which encodes the FLP site-specific recombinase, and was confirmed by replica plating on KB agar plates with and without kanamycin (House *et al.*, 2004). Finally, cells resistant to 10% sucrose were selected to identify those that lost pBH474.

***In planta* assay**

Bacterial cells were grown overnight in KB with appropriate antibiotics, washed and resuspended in 10 mM $MgCl_2$. For *in planta* growth assay, we resuspended *P. syringae* and *P. fluorescens* to an OD_{600} of 0.002 (1.0 x 10^6 cfu/ml). We used 1.0 ml syringes lacking needles to infiltrate bacterial suspensions into the abaxial side of leaves of 5 ~ 6 week-old plants. We used a number 2 cork borer to core four discs for each triplicate of each treatment at 0 and 3 days post inoculation (dpi). Leaf discs were ground to homogeneity in 10 mM $MgCl_2$, serially diluted, and plated on KB with appropriate antibiotics. Colonies were enumerated once visible. Experiments were repeated at least three times. For HR assays, cells were infiltrated in leaves at an OD_{600} of 0.2 (equivalent to 1.0 x 10^8 cfu/ml). Phenotypes were scored starting at six hours post inoculation (hpi) and examined up until disease symptoms were visible.

ACKNOWLEDGEMENTS

Cait Thireault contributed to this project by assisting with some of the many recombineering steps. This project also benefitted from the hard work of Stanley Lee and David Swader-Hines, who performed many of the PCR screens that determined successful integration of T3E-encoding clones into EtHAn and *Pph*1448a.

Table A2.1. Allocation of 28 *Pto*DC3000 T3E genes to clones 1-4.

	T3E gene*	Size (kb)
Clone 1	*hopAB2*	1.8
	hopI1	2.5
	hopAM1-1	1.0
	avrPto1	0.6
	avrE1	5.8
	shcM-hopM1	2.8
	shcE	0.4
Clone 2	*shcF2-hopF2-hopU1*	2.3
	shcO1-hopO1-1-hopT1-1	2.8
	hopD1	2.3
	hopG1	1.6
	hopC1	1.0
	hopH1	1.0
	hopAO1	1.5
Clone 3	*hopE1*	0.7
	hopAM1-2	0.9
	shcN1-hopN1	1.7
	hopK1	1.2
	hopP1	1.2
	hopY1	1.0
	hopAA1-1	1.8
	hopR1	6.1
Clone 4	*shcV1-hopV1*	1.8
	hopQ1-1	1.5
	hopAA1-2	1.7
	hopAF1	1.2
	hopX1	1.3
	shcA1-hopA1	1.7

*Full-length genes include *hrp*-box, cognate chaperone (if applicable), and other T3E genes present in operon, if any.

Figure A2.1. Construction of four clones carrying the T3E repertoire of *Pto*DC3000.
For each set of *Pto*DC3000 T3Es, a combination of recombinant and sticky-end PCR was used to amplify full-length effector-encoding genes or 100-bp regions from the 5' and 3' ends of the full-length effector-encoding genes respectively, adding restriction sites in the process. Amplified products were sequentially cloned into pBBR1-MCS1 via sequential restriction enzyme digests/ligations. Step A above represents one such set of T3Es, consisting of full-length *hopE1* and *hopAM1-2*, full-length *shcN1-hopN1* (encoding HopN1 and its cognate chaperone, ShcN1), and full-length *hopK1, hopP1,* and *hopY1*, as well as 100-bp regions from the 3' and 5' ends of the full-length T3E-encoding genes *hopAA1-1* and *hopR1*. In Step B, the full-length T3E-encoding genes *hopAA1-1* and *hopR1* were sequentially recombineered into pBBR1-MCS1. The plasmid is linearized at the restriction sites between the flanking regions and subsequently captures the full-length gene from the donor plasmid pDONR207. Step C represents the final recombineering step, in which a mini-Tn5 vector is linearized at an *Acc651* restriction site located between two 200-bp regions of homology to sequences flanking the set of full-length T3E-encoding genes in pBBR1-MCS1. The set of full length genes is captured by the mini-Tn5 vector and subsequently integrated into the genome of either EtHAn or *Pph*1448a via two-way mating and Tn5-mediated transposition. This process was repeated for four unique sets of T3Es, representing 29 T3Es of *Pto*DC3000.

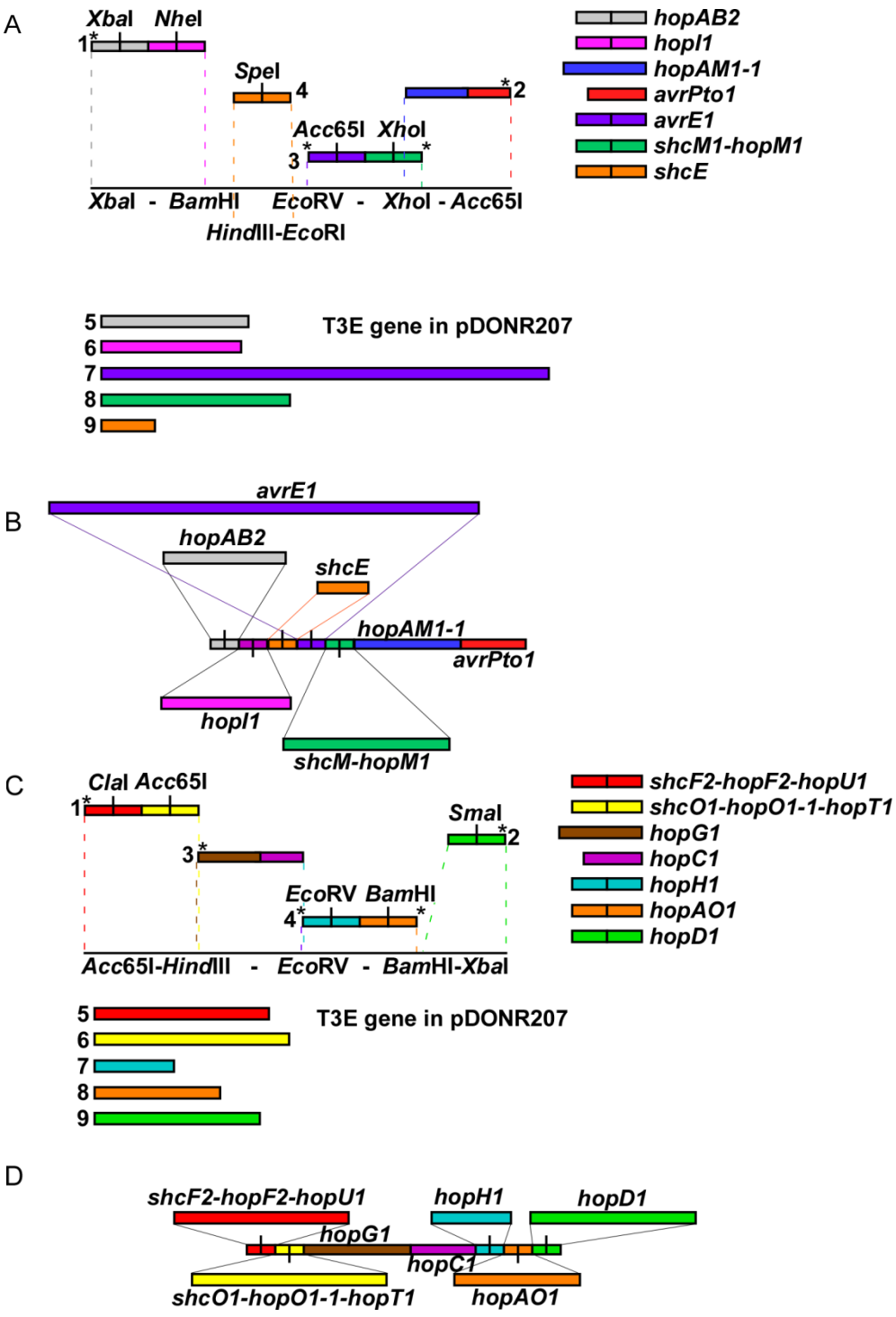


Figure A2.1. Construction of four clones carrying the T3E repertoire of *Pto*DC3000.

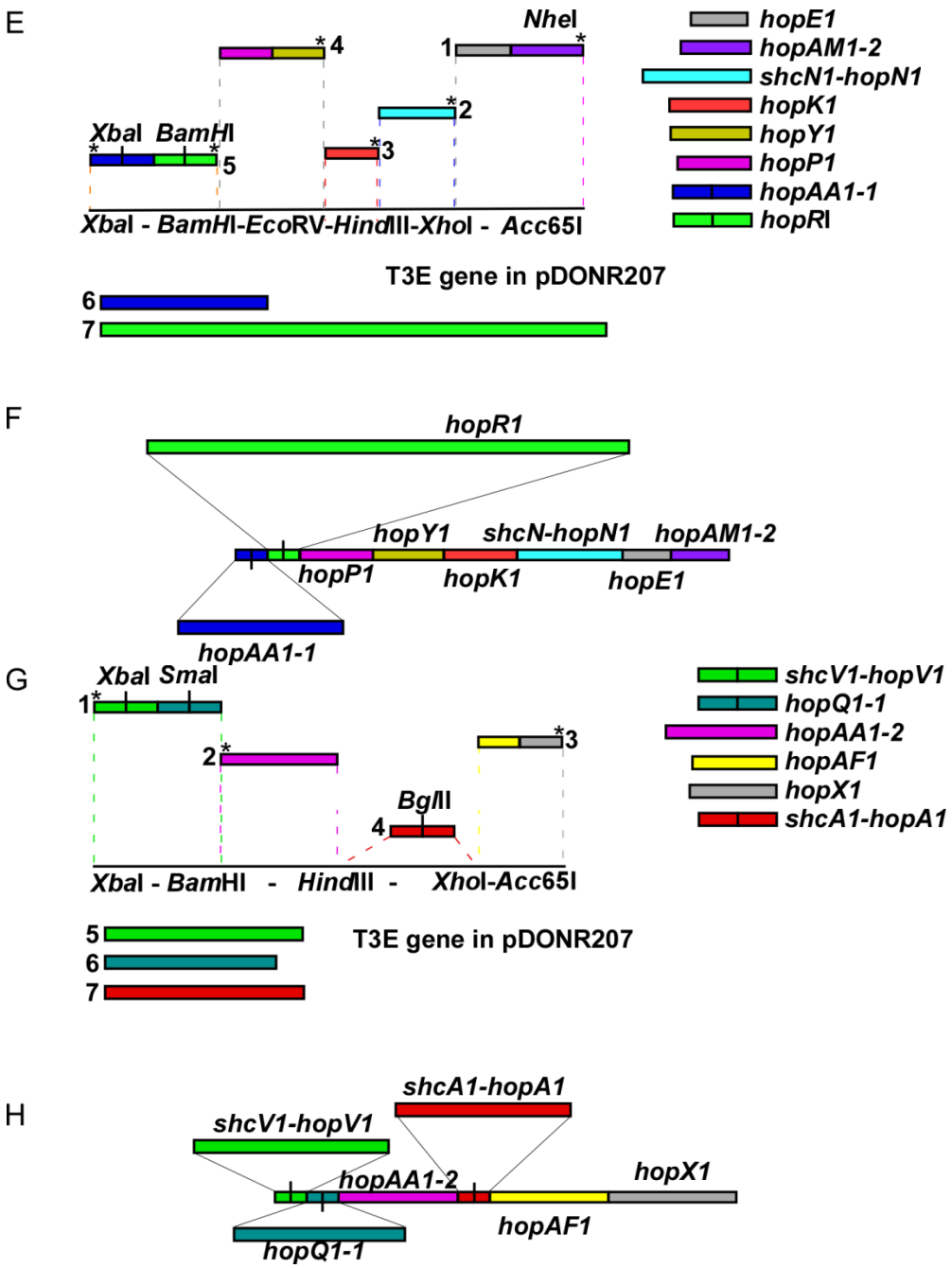


Figure A2.1. Construction of four clones carrying the T3E repertoire of *Pto*DC3000 (Continued).

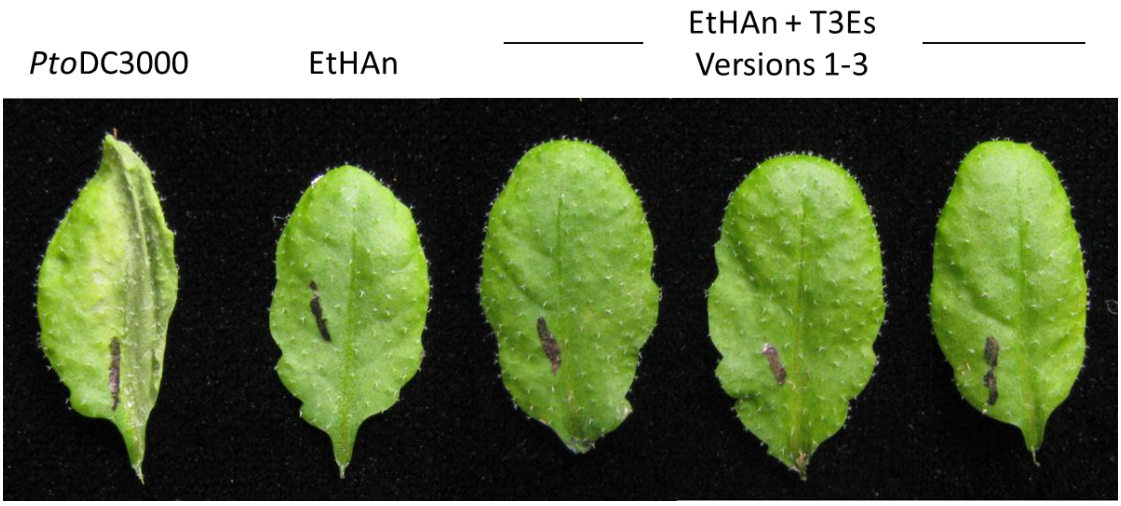


Figure A2.2. The *Pto*DC3000 T3E repertoire does not alter the EtHAn phenotype in Arabidopsis.

EtHAn encoding the *Pto*DC3000 T3E repertoire does not elicit a disease phenotype within 72 hpi. Each leaf pictured represents the phenotype observed in 100% of the leaves receiving the indicated treatment. This experiment was repeated twice with similar results.

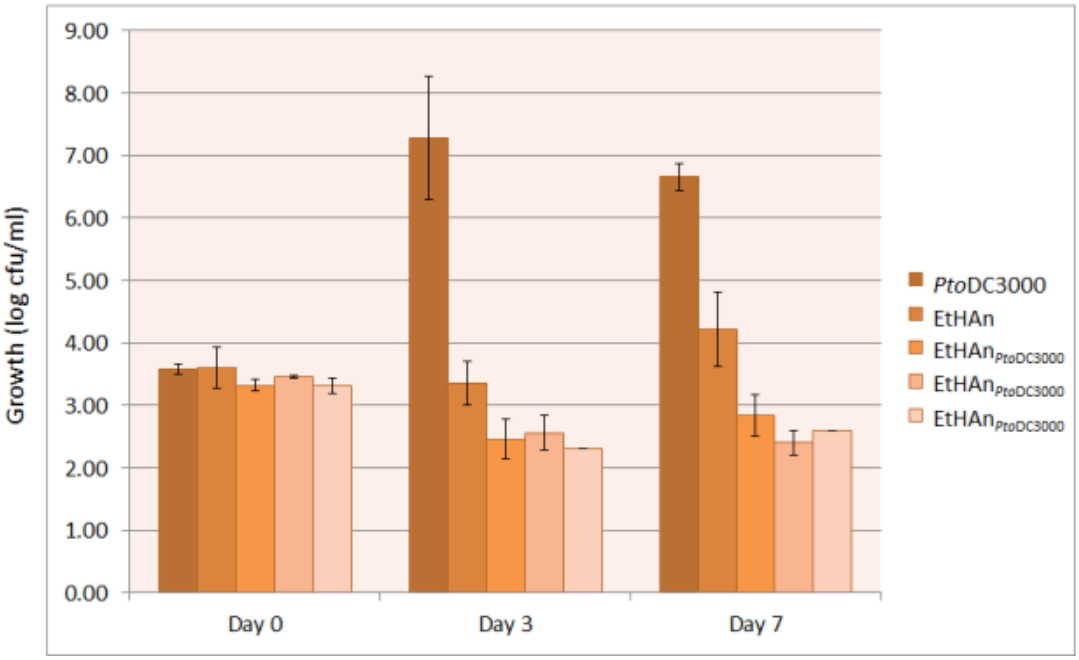


Figure A2.3. Enumeration growth *in planta* for EtHAn carrying *Pto*DC3000 T3E repertoire.

Wild-type *Pto*DC3000, EtHAn without T3Es, and three independently cloned variants of EtHAn encoding the *Pto*DC3000 repertoire were assessed for growth at 3 dpi and 7 dpi. Data points represent three biological replicates and standard error is shown.

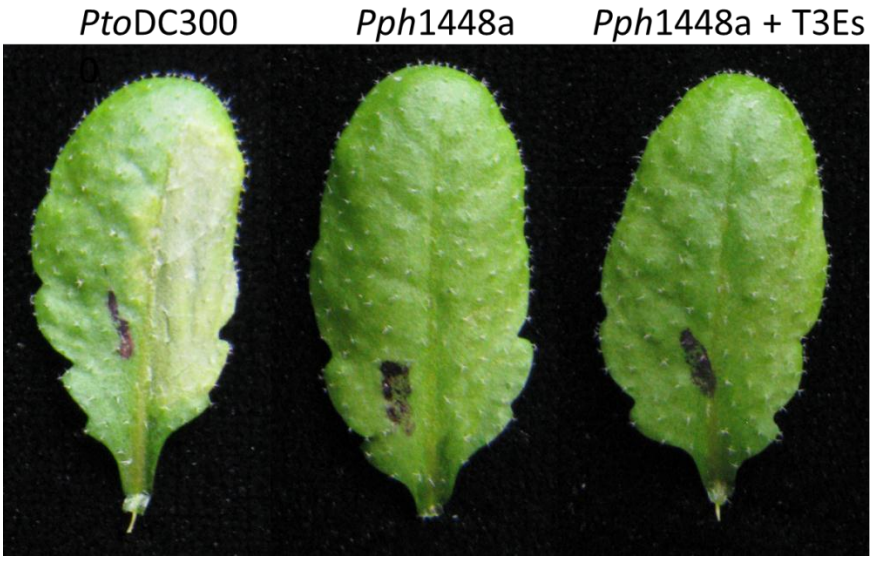


Figure A2.4. The *Pto*DC3000 T3E repertoire does not alter the EtHAn phenotype in Arabidopsis.
*Pph*1448a encoding the *Pto*DC3000 T3E repertoire does not elicit a disease phenotype within 72 hpi. Each leaf pictured represents the phenotype observed in 100% of the leaves receiving the indicated treatment. This experiment was repeated three times with similar results.

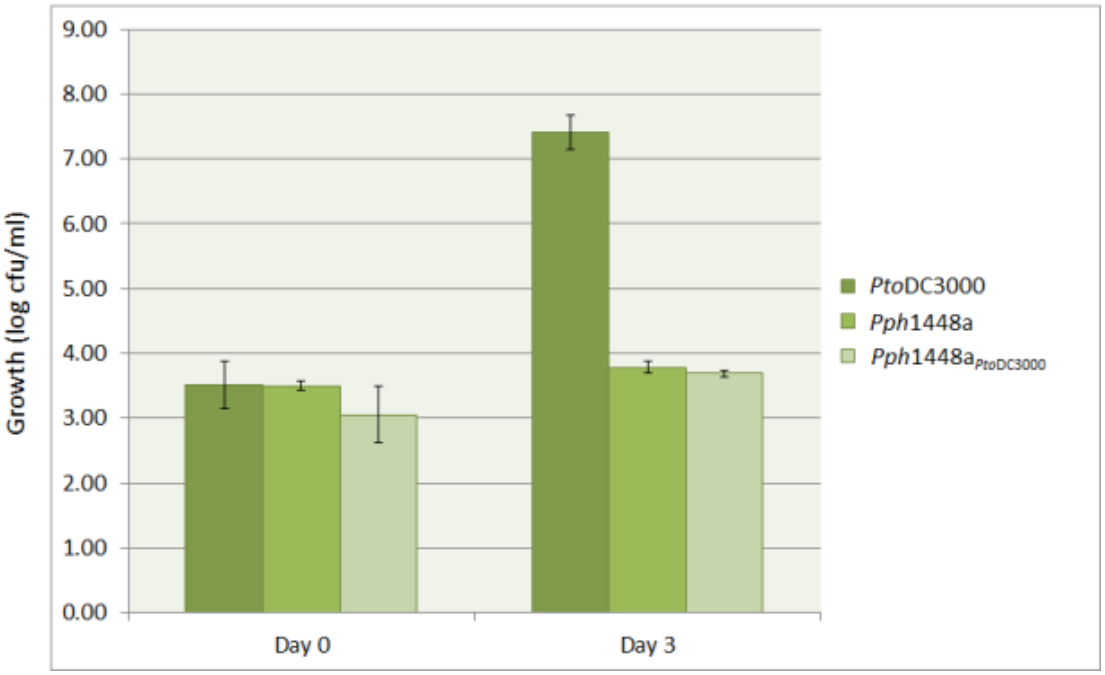


Figure A2.5. Enumeration growth *in planta* for EtHAn carrying *Pto*DC3000 T3E repertoire.

Wild-type *Pto*DC300, the T3SS-deficient *Pto*DC3000 mutant *ΔhrcC*, wild-type *Pph*1448a, and *Pph*1448a encoding the *Pto*DC3000 repertoire were assessed for growth at 3 dpi. Data points represent three biological replicates and standard error is shown.

REFERENCES CITED

An, C. and Mou, Z. (2011) Salicylic Acid and its Function in Plant ImmunityF. *Journal of Integrative Plant Biology*, **53**, 412-428.

Baltrus, D.A., Nishimura, M.T., Romanchuk, A., Chang, J.H., Mukhtar, M.S., Cherkis, K., Roach, J., Grant, S.R., Jones, C.D. and Dangl, J.L. (2011) Dynamic Evolution of Pathogenicity Revealed by Sequencing and Comparative Genomics of 19 <italic>Pseudomonas syringae</italic> Isolates. *PLoS Pathog*, **7**, e1002132.

Bao, Y., Lies, D.P., Fu, H. and Roberts, G.P. (1991) An improved Tn7-based system for the single-copy insertion of cloned genes into chromosomes of gram-negative bacteria. *Gene*, **109**, 167-168.

Chang, J.H., Urbach, J.M., Law, T.F., Arnold, L.W., Hu, A., Gombar, S., Grant, S.R., Ausubel, F.M. and Dangl, J.L. (2005) A high-throughput, near-saturating screen for type III effector genes from Pseudomonas syringae. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 2549-2554.

Cunnac, S., Occhialini, A., Barberis, P., Boucher, C. and Genin, S. (2004) Inventory and functional analysis of the large Hrp regulon in Ralstonia solanacearum: identification of novel effector proteins translocated to plant host cells through the type III secretion system. *Molecular Microbiology*, **53**, 115-128.

Datsenko, K.A. and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences*, **97**, 6640-6645.

de Lorenzo, V., Herrero, M., Jakubzik, U. and Timmis, K.N. (1990) Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. *Journal of Bacteriology*, **172**, 6568-6572.

Figurski, D.H. and Helinski, D.R. (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *Proceedings of the National Academy of Sciences*, **76**, 1648-1652.

Galan, J.E. and Wolf-Watz, H. (2006) Protein delivery into eukaryotic cells by type III secretion machines. *Nature*, **444**, 567-573.

Grant, S.R., Fisher, E.J., Chang, J.H., Mole, B.M. and Dangl, J.L. (2006) Subterfuge and Manipulation: Type III Effector Proteins of Phytopathogenic Bacteria. *Annual Review of Microbiology*, **60**, 425-449.

Ham, J.H., Kim, M.G., Lee, S.Y. and Mackey, D. (2007) Layered basal defenses underlie non-host resistance of Arabidopsis to Pseudomonas syringae pv. phaseolicola. *The Plant Journal*, **51**, 604-616.

House, B.L., Mortimer, M.W. and Kahn, M.L. (2004) New Recombination Methods for Sinorhizobium meliloti Genetics. *Applied and Environmental Microbiology*, **70**, 2806-2815.

Jamir, Y., Guo, M., Oh, H.-S., Petnicki-Ocwieja, T., Chen, S., Tang, X., Dickman, M.B., Collmer, A. and R. Alfano, J. (2004) Identification of Pseudomonas syringae type III effectors that can suppress programmed cell death in plants and yeast. *The Plant Journal*, **37**, 554-565.

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

Kovach, M.E., Phillips, R.W., Elzer, P.H., Roop II, R.M. and Peterson, K.M. (1994) pBBR1MCS: a broad-host-range cloning vector. *Biotechniques*, **16**, 800-802.

Lewis, J.D., Guttman, D.S. and Desveaux, D. (2009) The targeting of plant cellular systems by injected type III effector proteins. *Seminars in Cell & Developmental Biology*, **20**, 1055-1063.

Ma, W. (2011) Roles of Ca2+ and cyclic nucleotide gated channel in plant innate immunity. *Plant Science*, **181**, 342-346.

Mansfield, J.W. (2009) From bacterial avirulence genes to effector functions via the hrp delivery system: an overview of 25 years of progress in our understanding of plant innate immunity. *Molecular Plant Pathology*, **10**, 721-734.

Nürnberger, T., Brunner, F., Kemmerling, B. and Piater, L. (2004) Innate immunity in plants and animals: striking similarities and obvious differences. *Immunological Reviews*, **198**, 249-266.

Oh, C.-S. and Beer, S.V. (2005) Molecular genetics of Erwinia amylovora involved in the development of fire blight. *FEMS Microbiology Letters*, **253**, 185-192.

Pitzschke, A., Schikora, A. and Hirt, H. (2009) MAPK cascade signalling networks in plant defence. *Current Opinion in Plant Biology*, **12**, 421-426.

Postel, S. and Kemmerling, B. (2009) Plant systems for recognition of pathogen-associated molecular patterns. *Seminars in Cell & Developmental Biology*, **20**, 1025-1031.

Rafiqi, M., Bernoux, M., Ellis, J.G. and Dodds, P.N. (2009) In the trenches of plant pathogen recognition: Role of NB-LRR proteins. *Seminars in Cell & Developmental Biology*, **20**, 1017-1024.

Robert-Seilaniantz, A., Grant, M. and Jones, J.D.G. (2011) Hormone Crosstalk in Plant Disease and Defense: More Than Just JASMONATE-SALICYLATE Antagonism. *Annual Review of Phytopathology*, **49**, 317-343.

Sanabria, N.M., Huang, J.-C. and Dubery, I.A. (2010) Self/nonself perception in plants in innate immunity and defense. *Self Nonself.*, **1**, 40-54.

Schwessinger, B. and Zipfel, C. (2008) News from the frontline: recent insights into PAMP-triggered immunity in plants. *Current Opinion in Plant Biology*, **11**, 389-395.

Segonzac, C.c. and Zipfel, C. (2011) Activation of plant pattern-recognition receptors by bacteria. *Current Opinion in Microbiology*, **14**, 54-61.

Silby, M., Cerdeno-Tarraga, A., Vernikos, G., Giddens, S., Jackson, R., Preston, G., Zhang, X.-X., Moon, C., Gehrig, S., Godfrey, S., Knight, C., Malone, J., Robinson, Z., Spiers, A., Harris, S., Challis, G., Yaxley, A., Harris, D., Seeger, K., Murphy, L., Rutter, S., Squares, R., Quail, M., Saunders, E., Mavromatis, K., Brettin, T., Bentley, S., Hothersall, J., Stephens, E., Thomas, C., Parkhill, J., Levy, S., Rainey, P. and Thomson, N. (2009) Genomic and genetic analyses of diversity and plant interactions of Pseudomonas fluorescens. *Genome Biology*, **10**, R51.

Tena, G., Boudsocq, M. and Sheen, J. (2011) Protein kinase signaling networks in plant innate immunity. *Current Opinion in Plant Biology*, **14**, 519-529.

Thomas, W.J., Thireault, C.A., Kimbrel, J.A. and Chang, J.H. (2009) Recombineering and stable integration of the Pseudomonas syringae pv. syringae 61 hrp/hrc cluster into the genome of the soil bacterium Pseudomonas fluorescens Pf0-1. *The Plant Journal*, **60**, 919-928.

Zhou, J.-M. and Chai, J. (2008) Plant pathogenic bacterial type III effectors subdue host responses. *Current Opinion in Microbiology*, **11**, 179-185.

Zuber, S., Carruthers, F., Keel, C., Mattart, A., Blumer, C., Pessi, G., Gigot-Bonnefoy, C.c., Schnider-Keel, U., Heeb, S., Reimmann, C. and Haas, D. (2003) GacS Sensor Domains Pertinent to the Regulation of Exoproduct Formation and to the Biocontrol Potential of Pseudomonas fluorescens CHA0. *Molecular Plant-Microbe Interactions*, **16**, 634-644.