Title
Host Specificity and Virulence Mechanisms of Xanthomonas Type III Effector Proteins in Bacterial Spot Disease

Author
Schwartz, Allison Rose

Publication Date
2016

Peer reviewed|Thesis/dissertation

Host Specificity and Virulence Mechanisms of *Xanthomonas* Type III Effector Proteins in Bacterial Spot Disease

By

Allison Rose Schwartz

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Plant Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Brian J. Staskawicz, Chair
Professor Steven E. Lindow
Professor Suzanne M. Fleiszig

Fall 2016

Abstract

Host Specificity and Virulence Mechanisms of *Xanthomonas* Type III Effector Proteins in Bacterial Spot Disease

by

Allison Rose Schwartz

Doctor of Philosophy in Plant Biology

University of California, Berkeley

Professor Brian J. Staskawicz, Chair

Xanthomonas spp are the causative agents of bacterial spot disease on cultivated pepper, *Capsicum annuum*, and tomato, *Solanum lycopersicum*. Although pepper and tomato are closely related in the *Solanaceae*, four species of xanthomonads have differing host specificities and utilize unique virulence strategies between these two crops plants. A major factor differentiating these pathogens are the Type III Effector (T3E) proteins they deploy to overcome the plant's immune system and increase the host's susceptibility. The genetic diversity and composition of T3E repertoires in a large sampling of field strains have yet to be explored on a genomic scale, limiting our understanding of pathogen evolution in an agricultural setting. To this end, we sequenced the genomes of sixty-seven *Xanthomonas euvesicatoria* (Xe), *Xanthomonas perforans* (Xp), and *Xanthomonas gardneri* (Xg) strains isolated from diseased pepper and tomato fields in the southeastern and midwestern United States. T3E repertoires were computationally predicted for each strain and whole genomic phylogenies were employed to understand better the genetic relationship of strains in the collection. From this analysis we detected a division in the Xp population that supported a model whereby a host-range expansion of Xp field strains on pepper is due, in part, to a loss of the T3E AvrBsT. Xp-host compatibility was further studied with the observation that a double deletion of the T3Es AvrBsT and XopQ allowed a gain of host range for *Nicotiana benthamiana.* Additionally, a single deletion of XopQ expanded the host range of Xe to *N. benthamiana*, while Xg was a natural pathogen of *N. benthamiana*. Extensive sampling of field strains and an improved understanding of effector content will aid in efforts to design plant disease resistance strategies targeted against highly conserved effectors.

Xg has emerged recently as the dominant tomato pathogen in parts of the United States and South America. It is responsible for severe crop losses and causes spotting on fruits. Furthermore, Xg appears to be spreading globally. In its repertoire of Type III effectors, Xg possesses a single Transcription Activator Like (TAL) effector protein, AvrHah1, which has previously been shown to confer enhanced water soaked lesions in pepper. TAL effectors act as transcription factors that manipulate expression of target host genes to increase host susceptibility. We investigated the molecular mechanism of AvrHah1-dependent water

soaking and the effects of water soaking on enhancing disease severity in tomato. We observed that water from outside the leaf was drawn into the apoplast in Xg-, but not XgΔAvrHah1-, infected tomato, and that water soaking can serve as a mechanism to "ferry" new bacteria into the apoplast. Additionally, AvrHah1 increased the bacterial population present on the surface of diseased tomato leaves. Comparing the transcriptomes of tomato infected with Xg wt vs XgΔAvrHah1 revealed that thousands of genes were differentially upregulated in the presence of AvrHah1. We identified two highly upregulated basic Helix Loop Helix (bHLH) transcription factors with predicted Effector Binding Elements (EBEs) as direct targets of AvrHah1. We mined our RNA-seq data for genes that were highly upregulated but without EBEs and identified two pectin modification genes, a pectate lyase and pectinesterase, which are expressed in response to the bHLH transcription factors and are therefore indirect targets of AvrHah1. Importantly, designer TAL effectors (dTALEs) for the bHLH transcription factors and the pectate lyase complement water soaking in XgΔAvrHah1. By modifying the plant cell wall to enhance water uptake and increase tissue damage, AvrHah1 may improve bacterial dispersal from the apoplast and thereby enhance disease transmission. Understanding lesion development may improve the design of disease tolerance in crops by reducing symptom development and overall pathogen transmission.

Acknowledgements

I would like to acknowledge and thank all the people who supported me during my time as a doctoral student at UC Berkeley. Firstly, I'd like to thank Brian Staskawicz for being a supportive Ph.D. advisor. Brian gave me the opportunity to pursue my own scientific questions while providing excellent guidance and for that I am extremely grateful. I'd like to thank Doug Dahlbeck for helping me become a better molecular biologist, and for his knowledge, tools, and resources that have helped everyone in the lab. I'd like to thank Steve Lindow for being a great sounding board and for his enthusiasm and support throughout my rotation in his lab, my qualifying exam, and my thesis committee meetings.

I would also like to thank and acknowledge my collaborators for their contributions to this research. Thank you to Neha Potnis for population assays, Sujan Timilsina, Rebecca Bart, Erica Goss, Mark Wilson, José Patané, and Joaquim Martins Jr. for their expertise in whole genome phylogenies, Sally Miller, Gerald Minsavage, Gary Vallad, and David Ritchie for bacterial strains, Robert Morbitzer for dTALE construction, and Thomas Lahaye, João Setubal, Jeffrey Jones, and Frank White for reading and commenting on manuscripts.

My first day in the Staskawicz lab, Brian suggested treating graduate school like a marathon rather than a sprint. I have many friends to thank that have cheered me on during this experience and they have made all the difference to help me stay motivated and finish strong. I had many wonderful labmates in the Staskawicz lab, but I would especially like to acknowledge Megan and Adam, who were great friends and inspiring scientists. To my fellow Plant Biology cohort Anne, Jason, Mike, and Ping-hung, thank you so much for all the coffee dates, game nights, and science help. To my UCLA ladies Christine, Shannon, Tina, and Ariel, you are amazing an inspiring women and I am so grateful for your friendship. To my writing buddies Jeanne and Ioana, you came into my life at just the right moment. I am honored and grateful to have shared the dissertation writing experience together. Thanks for the dance parties, the late nights, and the wise mantra, "You do you!"

I would like to specially acknowledge my undergraduate research advisor at UCLA, Ann Hirsch. Ann's excellent teaching and enthusiasm for plant-microbe interactions and dedication to undergraduate researched launched me into this field, and I strive in my own scientific career to emulate her excellence. I would also like to thank Nancy and Annette, who I got to know when they were post-doctoral researchers in the Hirsch lab. They showed me the joy of science and their generous, graceful, and kind spirits have inspired me to grow as a person and as a scientist. MEEP!

I would like to thank my roommate, John, and his mother, Seldy, for welcoming me into their lovely family. I am extremely thankful to have both of them in my life.

I could not have done this work without my family. I would like to thank my sister, Laurel, for always being there for me and for my brother, Ethan, who reminds me of the importance of being brave and following your passions. This dissertation is dedicated to my parents, Rob and Leslie, because their unconditional love and support has allowed me to be the best version of myself. Thanks Mom and Dad, I love you guys.

Table of Contents

List of Figures

1. Introduction

2. Phylogenomics and Type III effector predictions of *Xanthomonas* field strains

3. Two Type III effectors define host range of Xe and Xp, but not Xg

4. Characterization of water soaking from AvrHah1, a TAL effector from Xg

5. Direct and indirect targets of AvrHah1 identified using RNA-seq have to ability to complement water soaking in XgΔAvrHah1

6. Differential recognition of AvrBs3 and AvrHah1 by the tomato R protein Bs4

7. Investigation of delivery of plant "effectors" via the Type III Secretion System

List of Tables

List of Abbreviations

20R	Pepper with Bs2 R gene backcrossed into ECW
30R	Pepper with Bs3 R gene backcrossed into ECW
aa	Amino acid
AD	Activation Domain
Avr	Avirulence
bHLH	Basic Helix Loop Helix
DBD	DNA Binding Domain
DPI	Days Post Inoculation
dTALE	Designer TAL effector
EBE	Effector Binding Element
ECW	Early CalWonder
ETI	Effector Triggered Immunity
HGT	Horizontal Gene Transfer
HPI	Hours Post Inoculation
HR	Hypersensitive Response
ML	Maximum Likelihood
MLSA	Multilocus Sequencing Analysis
PAMP	Pathogen Associated Molecular Patterns
PTI	PAMP Triggered Immunity
R	Resistance
RL	Reflected Light
RPKM	Reads Per Kilobase of transcript per Million mapped reads
RVD	Repeat Variable Diresidue
S	Susceptibility
SNP	Single Nucleotide Polymorphism
T3SS	Type III Secretion System

TAL	Transcription Activator Like
TL	Transmitted Light
Wt	Wild type
Xe	*Xanthomonas euvesicatoria*
Xg	*Xanthomonas gardneri*
Xop	*Xanthomonas* outer protein
Xp	*Xanthomonas perforans*

1. Introduction

Species of *Xanthomonas* cause bacterial spot disease on cultivated pepper (*Capsicum annuum*) and tomato *(Solanum lycopersicum)* and are the most devastating to crops grown in warm, humid climates such as in the southeastern and midwestern United States. Once considered a single species, *Xanthomonas vesicatoria* infecting pepper and tomato has been reclassified several times (1-3), but was most recently separated into four distinct species based on multilocus sequence typing: *X. euvesicatoria* (Xe), *X. vesicatoria* (Xv), *X. perforans* (Xp), and *X. gardneri* (Xg) (4). While *Xe, Xg, and Xv* infect both pepper and tomato, Xp has been limited to tomato (5). The diversity of these four species and their convergence onto tomato and pepper as hosts make them an interesting pathosystem to study common virulence mechanisms and factors contributing to host specificity (6).

Although the four pathogens are present and destructive on a global scale (7, 8), the history and distribution of Xe, Xp, and Xg has changed dramatically in the United States, particularly with the emergence of Xp as the dominant tomato pathogen over Xe in Florida beginning in the early 1990's (5, 9-12) and Xg as a major tomato pathogen in Ohio and Michigan beginning in 2009 (13). Outbreaks of Xv have not been reported in the United States (8).

Xe, Xp, Xg, and Xv are xanthomonads with a typical yellow, mucoid colony appearance (Fig. 1-1). Outbreaks of bacterial spot are worsened during periods of warmer temperatures and high humidity (100). Once on the leaf surface, these foliar pathogens access the apoplast through wounds or natural openings, such as stomata or hydathodes. Lesions can appear as initially water-soaked but progress into a necrotic brown or "shot-through" appearance. Defoliation reduces fruit yield, and lesions on fruits further reduce marketability (5). Although disease lesions will appear on all aboveground organs, most of the research on xanthomonads and their hosts has focused on infected leaves. Management of bacterial spot is complicated by the ability of xanthomonads to survive as epiphytes on seeds, non-host plants, and decaying plant material, which serve as future inoculum for neighboring plants (100).

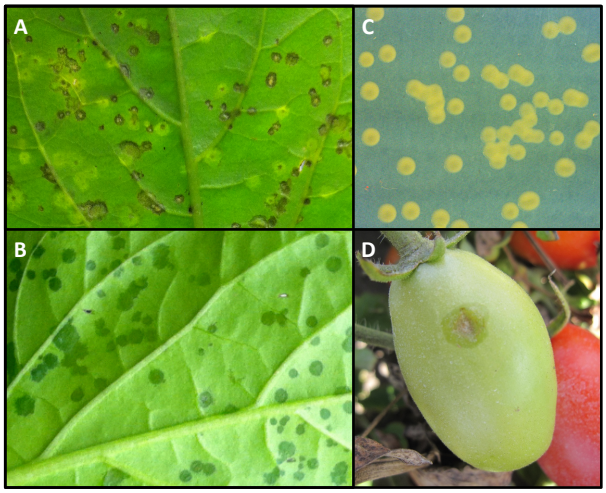


Figure 1-1. Bacterial spot caused by *Xanthomonas*. (A) Pepper ECW infected with Xg shows water soaked and nectrotic lesions on the adaxial (B) abaxial leaf. (C) Mucoid and yellow appearance of Xg. (D) Water soaked lesion from Xg on tomato fruit (photo credit Sally Miller).

Phylogenetic analyses have revealed a close evolutionary relationship between Xe and Xp in comparison to Xg and Xv, which are separated evolutionarily by xanthomonads infecting diverse crops such as rice and citrus (14-17) (Fig. 1-2). A major factor contributing to the virulence and host specificity of these pathogens is the repertoire of effectors secreted into the host plant cell via the Type III Secretion System (18). Xanthomonads have evolved effectors with diverse mechanisms to promote virulence and suppress the plant immune response, remarkable in that they have evolved to adopt processes specific to eukaryotes (19, 20). The recognition of specific effector proteins by specific cognate resistance (R) proteins leads to defense responses that have been termed Effector Triggered Immunity (ETI), which is accompanied by localized cell death, an associated tissue collapse known as the Hypersensitive Response (HR) at the site of infection, and limited spread of the pathogen (21). Effector genes are thus under strong selective pressure to acquire mutations that allow their protein products to evade recognition by plant R-proteins, yet maintain a virulence effect (18). Several Type III effectors are conserved across multiple species and referred to here as core effectors. The identification or design of R-proteins that can recognize core effectors may provide durable and broad-spectrum resistance. Thus, the identification of core effectors from a representative sampling of field isolates would provide valuable information to inform future genetic resistance strategies. An additional variable set of effectors may provide specialization to specific hosts and cultivars (22).

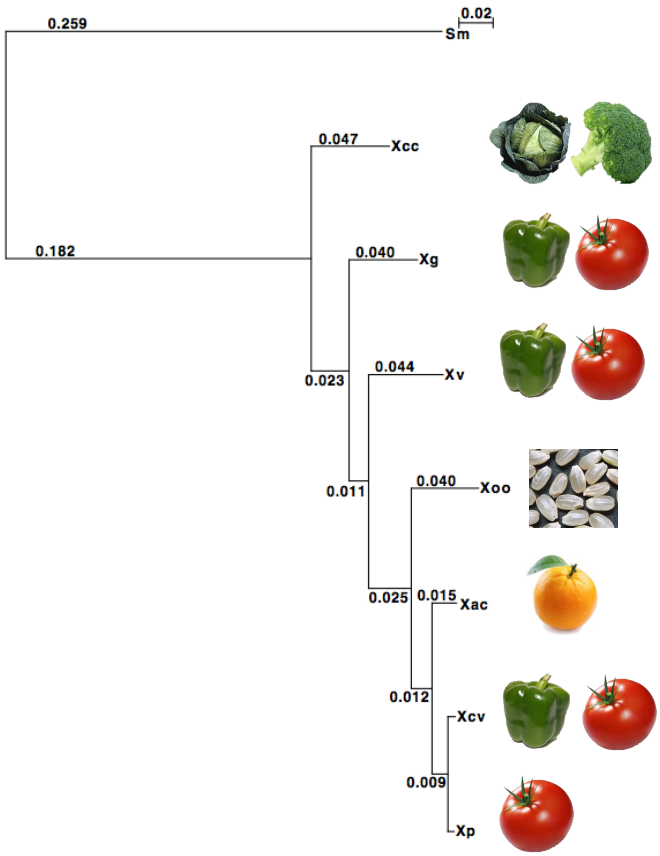


Figure 1-2. Phylogenetic relationships of xanthomonads infecting pepper and tomato. Xg, Xv, and Xcv (now Xe) infect both pepper and tomato and Xp infects tomato. Xe and Xp are separated from Xg and Xv by *Xanthomonas oryzae* pv. *oryzae* (Xoo), a pathogen of rice and *Xanthomonas axonopodis* pv. *citri* (Xac), a pathogen of citrus. Modified from (6).

The deployment of R proteins in crops that can respond to core effectors is an effective disease resistance strategy, depending on the evolutionary stability of the targeted cognate effector (23). For example, introgression of the wild pepper (*Capsium chacoense*) *Bs2* gene into cultivated pepper provided resistance against multiple species of xanthomonads that infect pepper because of the pervasiveness of the core effector AvrBs2, which is required for full virulence and is widespread across xanthomonads infecting diverse hosts (101). However, strains

overcoming Bs2 resistance, due to mutations in AvrBs2 that evade Bs2 recognition, were reported in the field only three years after commercialization of Bs2 pepper (5). Thus, the long-term durability of future disease resistance strategies must include the pyramiding of multiple genetic resistances in order to slow the emergence of pathogenic strains (23, 24). Despite successful field trials of transgenic *Bs2* tomato there is no genetic resistance currently used in commercial tomato farming (12). Due to the lack of genetic resistance, streptomycin and copper-containing antimicrobial sprays have been the main means of defense against tomato pathogens. Expectedly, the development of bacterial resistance to these antimicrobials has made these treatments ineffectual (5).

The severe economic consequences of tomato bacterial spot have necessitated the search for disease-resistance in wild species (that could be crossed with commercial agronomic cultivars) and more recently targets for DNA editing resistance strategies. Before designing new genetic resistance strategies for crops, an understanding of the pathogens the plants will encounter is first required. Comparative genomics of the reference strains Xe85-10 (25), Xp91-118, Xg 101 and Xv 1111 provided the first insights into the shared and unique virulence factors of these pepper and tomato pathogens (6). However, because xanthomonads display relatively high genome plasticity, a more comprehensive understanding of the genetic diversity of pepper and tomato pathogens, with specific emphasis on effectors, is necessary for designing informed disease resistance strategies for agricultural areas afflicted by bacterial spot disease (6, 8, 25).

A comparative genomic analysis considering many strains from a given geographic region over time will provide a representative view of the effectors present in the regional bacterial population and add insight into the evolutionary trends of effectors, and thus their potential usefulness as targets for R-gene mediated resistance strategies. To this end we sequenced the genomes of 32 Xp, 25 Xe, and 10 Xg field strains that were collected from diseased peppers and tomatoes in the southeastern and midwestern United States. We describe the genetic diversity within and between species using core protein-coding genome phylogeny and whole genome single nucleotide polymorphism (SNP) analysis and present the computationally predicted Type III effector repertoires of strains in our collection. We identified two new core effectors commonly shared between Xp, Xe, and Xg and show the many differences in effector sequences between the reference and field strains. From our phylogenomic approach we identified a split in the Xp population that correlated with host specificity on tomato and the presence of the effector AvrBsT. We further explored host specificity of AvrBsT with XopQ on *Nicotiana benthamiana* for Xe and Xp. We discovered that Xg is pathogenic on *Nicotiana benthamiana*, explained by an un-recognized allele of XopQ.

Xg has emerged relatively recently as the dominant tomato pathogen in the midwestern United States and causes significant spotting on fruits (13). The Xg strains sequenced in our collection were strains from Ohio and Michigan, collected between 2010 and 2012 and represent an opportunity to study the virulence of an emerging pathogen. Xg is also responsible for severe crop losses in Brazil (26) and appears to be spreading globally (8). Among its repertoire of Type III effectors, Xg possesses a single Transcription Activator Like (TAL) effector protein, AvrHah1, which has been shown to confer enhanced water soaked lesions in pepper (27). TAL effectors are secreted into host plant cells via the bacterial Type III Secretion System and are delivered into the nucleus, where they activate the expression of target host genes to increase host susceptibility (28). The TAL effector

binding specificity for host target DNA depends on the amino acid sequence of the central DNA Binding Domain (DBD), which is composed of several repeats of 34-35 amino acids, nearly identical except for the 12th and 13th amino acids of each repeat, termed the Repeat Variable Di-residue (RVD) (29). Each RVD confers binding specificity to a particular nucleotide, and in combination the targeted host sequence, termed the "Effector Binding Element (EBE), can be predicted." (29, 30). Once bound to a DNA target, the acidic Activation Domain (AD) of the TAL effector recruits the host's transcriptional machinery to activate gene expression (28).

Depending on the identities of the TAL effector RVDs and the corresponding EBEs in plant hosts, TAL effectors may target a single gene or multiple genes. If the plant gene

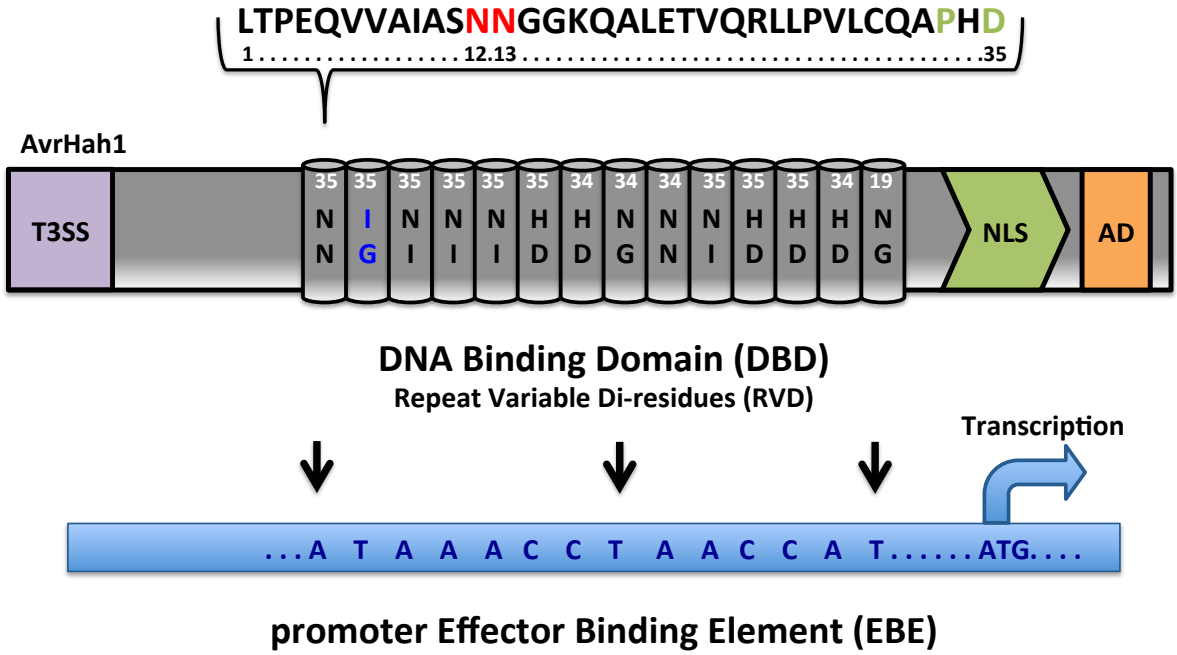


Figure 1-3. Structure of AvrHah1, a Transcription Activator Like (TAL) effector
AvrHah1 is a 1,039aa protein with a N-terminal Type III Secretion Signal (T3SS), central DNA Binding Domain (DBD), and C-terminal Nuclear Localization Signal (NLS) and transcriptional acidic Activation Domain (AD). The DBD is composed of 13.5 nearly identical repeats of either 35 or 34 amino acids (in white text). 35aa repeats have an extra Proline and variable terminal amino acid (in green). The 12th and 13th amino acid of each repeat is termed the Repeat Variable Di-residue (RVD) and confer binding to a specific nucleotide. "IG" (in blue) is a RVD that has only been reported in AvrHah1.

target of the TAL effector promotes pathogen virulence, the gene is considered a susceptibility (S) gene. Identifying and characterizing S genes reveals pathogen strategies and is useful in the design of more disease resistant plants through the removal of EBEs from the host gene promoter (31). Designer TAL effectors, or dTALEs, assembled with repeat modules designed to bind to a specific nucleotide sequence, can be used probe the function of the plant S gene (32). In the case of TAL effectors that activate multiple host genes, it becomes increasingly challenging to identify the *bofa fide* S gene(s) (33). Determining whether a target is a *bona fide* S gene has largely relied on the definition of virulence applied to effector biology, which is a significant growth benefit *in planta*. However, *in planta* growth assays do not take into account key events in the pathogen life cycle, such as emergence to the plant surface or transmission to neighboring plants.

Several examples connecting water-soaked lesion development and TAL effector S gene targets have been reported in diverse plant-xanthomonad pairs. In rice, Tal2g from *X. oryzae pv. oryzicoa* (Xoo) activates expression of *OsSULTR3;6*, which encodes a sulfate transporter (34). Mutations in Tal2g reduced lesion expansion and Xoo surface population, but not *in planta* growth, and dTALEs activating expression of *OsSULTR3;6* restored lesion expansion and surface growth to wild type levels (34). *CsLOB1*, a member of the lateral organ boundaries family from citrus, is activated by the PthA family of TAL effectors from xanthomonad pathogens of citrus. Loss of PthA reduces *in planta* bacterial growth and pustule formation (35, 36). TAL20 from the vascular cassava pathogen *X. axonopodis pv. manihotis* (Xam) activates expression of *MeSWEET10a*, which encodes a sugar transporter (37). dTALEs activating *MeSWEET10a* complemented the reduction in water soaking and midvein bacterial growth displayed by a Xam *TAL20* deletion strain in cassava (37). Avrb6 in *X. campestris pv. malvacearum* correlates with increased water soaking and bacterial surface population in cotton (38). AvrBs3 activates the pepper basic Helix Loop Helix (bHLH) transcription factor UPA20 which induces cell hypertrophy (39). Loss of AvrBs3 has been demonstrated to incur a bacterial fitness cost in the field (40).

Previous work has demonstrated that while AvrHah1 does not promote apoplastic growth, it does contribute to ion leakage and water soaking development in pepper (27). We wondered if we could use AvrHah1 as a tool to understand how water soaking develops in plant hosts. We observed that AvrHah1 enables the absorption of water into the apoplast of Xg-infected leaves, conferring a dark, water soaked appearance. The AvrHah1-mediated intake of water can be observed in real time and can also be measured quantitatively by collection and weighting of apoplastic fluid. Furthermore, bacterial cells can be ferried into the apoplast during water soaking. Interestingly, AvrHah1 is able to evade recognition by the tomato Bs4 R protein (in contrast to AvrBs3). We performed comparative RNA sequencing analysis of tomato leaves infected with Xg wild type or XgΔAvrHah1 and identified two bHLH transcription factors that were highly upregulated in the presence of AvrHah1. We show that two pectin modification genes—a pectate lyase and a pectinesterase—are downstream targets of both bHLH transcription factors and are therefore indirect targets of AvrHah1. We constructed dTALEs targeting the bHLH transcription factors and the pectin modification genes for gene activation and show that dTALEs activating the bHLH transcription factors and the pectate lyase complement the water soaking defect of XgΔAvrHah1. To our knowledge, this is the first report of virulence activity from an indirect TAL effector target.

2. Phylogenomics and Type III effector predictions of *Xanthomonas* field strains

Background

X. euvesicatoria (Xe), *X. perforans* (Xp), and *X. gardneri* (Xg) are three distinct species of xanthomonads that cause disease on pepper and/or tomato. Prior to this work a single reference genome existed for these species: Xe85-10, Xp91-118, and Xg 101, isolated in 1985, 1991, and 1953, respectively. These genomes provided important information about the genetics of these pathogens, particularly their shared and unique virulence mechanisms. However, a single reference genome of a type strain does not represent the current genetic diversity of strains present in the field in the United States. This is especially true of Xg 101, which was isolated over 60 years ago in Eastern Europe. Understanding the genetic diversity and evolutionary trajectories of effectors is particularly important when planning R-protein strategies for crop disease protection. To this end we sequenced the genomes of strains of Xe, Xp, and Xg isolated by our collaborators over the past 18 years. Besides temporal diversity, strains from this collection cover two major tomato and pepper growing regions in the United States. We used a phylogenomics approach incorporating core genes shared between Xe, Xp, and Xg as well as whole genome SNP analysis in order to study the genetic diversity within our collection of field strains. We detected a split in the Xp population not detected with previous MLST approaches. We computationally predicted the Type III effectors using an in-house algorithm, as previously described using a database of known effector sequences (41). We uncovered several differences in effector repertoires (including presence/absence and different alleles) between the reference and field strains and identified new core effectors shared between Xe, Xp, and Xg.

This research section shows the utility of sequencing multiple field strains from a given geographic region and the degree of genetic diversity between strains in the field. The spread of agricultural pathogens into new niches, either by increasing global movement of food crops or the emergence of new niches from climate change, makes the continued genomic surveillance of agricultural pathogens a top priority for food security and resistance strategies.

Results

Collection, genome sequencing, and genome assembly of *Xanthomonas* field strains

Xanthomonas field strains were collected by our collaborators and represent 18 years of evolution in the field in the major tomato growing regions in the Southeastern and Midwestern United States. Year, host, location, and isolation source are described in Table 1. Single colonies were isolated from field samples and stored at -80^{o}C. DNA was extracted from single colony-grown cultures using a modified CTAB method and used for in-house sequencing library construction as previously described with adaptors and indices for Illumina sequencing (41). Average *de novo* assembly statistics for 32 Xp, 25 Xe and 10 Xg are presented in Table 2. Assembly statistics for each strain can be found in Table S1. Draft genome sequences have been deposited in the National Center for Biotechnology (Table S1).

Core genome phylogenetic analysis identifies a division in the *Xp* population

The core genome for all three species was identified by sequence similarity, yielding 1,152 protein-coding gene families, of which 1,017 were considered bona fide orthologs; 135 families were discarded as spurious alignments by the program Guidance. The 1,017 families were concatenated, yielding a supermatrix of 916,326 sites. The best partitioning scheme chosen was by codon position in which first, second and third positions are set as separated partitions. The best evolutionary models for each partition were respectively GTR+I+G for the first and second partitions, and TVM+I+G for the third partition.
The Maximum Likelihood (ML) phylogeny based on core genome orthologs displays Xe, Xp, and Xg behaving as separate monophyletic groups (Fig. 2-1). Our results mirrored previous studies, Xe and Xp being closely related, and Xg more distant phylogenetically, with all three species forming monophyletic groups. For Xp strains, this analysis showed a division, which we define here as Group 1—further divided into Group 1A and 1B—and Group 2. Group 1A comprises eleven strains (out of sixteen) from 2012 that form a monophyletic clade (branches in purple). Other strains belonging to Group 1 are defined here as Group 1B (branches in orange), which includes the reference strain Xp91-118, Xp4B (isolated in 1998), and six strains isolated in 2006. Group 1B does not contain any strains isolated in 2012. We define fourteen strains as Group 2 (branches in green) which includes five strains from 2006, the single strain from 2010, five strains from 2012, and all three 2013 strains.

Whole genome SNP analysis resolves genetic differences among closely related strains

A total of 225,284 SNPs were identified between the Xe, Xg and Xp genomes compared to the reference Xac306, ranging from 22,105 (Xg 164) to 142,272 (GEV1063). Average SNPs (± standard deviations) between Xac306 and Xe, Xp and Xg field strains are 128376±3024, 136673±3402, and 30,462±8015, respectively. SNPs were concatenated and used to build a combined species ML tree (Fig. 2-2). We note that differences in sequencing technology used, genome coverage and large deletions or insertions could potentially skew this analysis and therefore conclusions about branch length between the different species should be avoided. The Xp Group 1A clade is retained in the ML SNP phylogeny (branches marked in purple). However Group 2 (green branches) is interrupted by Group 1B strains (orange branches).

Effector predictions for *Xanthomonas* field strains identifies differences in effector content compared to reference genomes

Type III effector repertoires from Xe, Xp, and Xg field strains were compared to the appropriate reference strains Xe85-10, Xp91-118 and Xg 101 in order to determine if effector repertoires differed between strains with respect to the presence or absence of whole effectors, mutations rendering effectors inactive, or alternate alleles of effectors.

Xe effector predictions

Several differences were found in the effector content of Xe field strains compared to the reference Xe85-10 (Table 3). Firstly, Xe85-10 does not have the effector XopAE, which is a translational fusion of the *hrp* cluster members *hpaG* and *hpaF* as seen in Xp91-118 (6). Similar to Xe85-10, field strains isolated before 1997 have separate *hpaG* and *hpaF* genes, whereas *Xe* field strains isolated after 1997 possess the predicted *hpaG*/*hpaF* translational fusion XopAE. Secondly, strains collected after 1997 possess a XopAF-like effector. The effector has 31% amino acid identity to XopAF of Xp91-118, 80% amino acid identity to *X. fuscans* XopAF (WP_022560489.1) and is identical to an effector of *X. citri* pv. *citri* (WP_015472934.1) except for an in-frame internal 12 amino acid deletion. Similarly, the Xe strains isolated after 1997 possess XopE3, which shares 97% amino acid identity with XopE3 from *X. arboricola* pv. *pruni* (WP_014125894.1). All field strains of Xe but one lack XopG, which is carried by the reference strain Xe85-10. A predicted protein-tyrosine phosphatase (abbreviated PTP) was detected in Xe075 that is not present in any other Xe strains. Twelve effectors present in all Xe strains isolated between 1985 and 2012 have no nucleotide polymorphisms. Xe field strains in our collection isolated after 1997 did not contain polymorphisms in *xopAA*, *xopF1*, *xopN*, and *xopO*. Except for Xe85-10 and Xe075, all Xe strains have identical sequences for effectors *xopAI*, *xopQ* and *xopV*.

The neighbor-joining tree of the effector alleles displays a grouping of the seven Xe strains isolated before 1997, and a clade of eleven strains with nearly identical allele profiles isolated from 2004-2012 (Fig. 2-3A). Although Xe111 and Xe112 group with the

clade of eleven strains and were isolated in Georgia in 2004, two other Georgia 2004 strains, Xe109 and Xe110, are separated from this clade due to differences in *avrBs2*, *xopE2*, and *xopO*. Interestingly, Xe082 was isolated in 1998 but has an effector allele profile similar to the eleven-member clade made up of strains isolated between 2004-2012.

Xp effector predictions

A shift in pathogen populations from tomato race 3 to tomato race 4 has been observed in Florida (12). All the strains sequenced here (with isolation years spanning from 1998 to 2013 in Florida) are tomato race 4 strains and contain null mutations in the *xopAF*/*avrXv3* gene of the reference strain Xp91-118 (Table 4). All strains possess XopJ4/AvrXv4 with the exception of the pepper strain Xp2010. Another effector, AvrBsT, which has been associated with hypersensitive response (HR) on pepper (42), has not been previously reported in Xp. AvrBsT is also present in nine strains (out of eleven) that were isolated in 2006, in all sixteen strains collected in 2012, and in one of the three strains collected in 2013. Interestingly, strain Xp17-12 (isolated in 2006) contains two effectors, XopF2 and XopV, that have sequences identical to the corresponding Xe85-10 effector sequence (Table 3). Effectors XopD and XopAD exhibit different alleles in the strains isolated in 2012. All strains have XopE2, which was absent in the reference strain Xp91-118. XopE2 is also present in all Xe and Xg field and reference strains. A subset of the Xp 2006 population have XopE4, which had been reported only in *X. fuscans* pv. *aurantifolii* (Moreira *et al.*, 2010). However, XopE4 is not present in any strains from 2010, 2012, or 2013. Interestingly, strains belonging to Xp Group 2 possess a XopQ identical to the allele from Xe85-10. The neighbor-joining tree based on effector alleles shows the conservation of Group 1A, but Group 1B and Group 2 strains were intertwined (Fig. 2-3B).

Xg effector predictions

The first reports of tomato bacterial spot caused by Xg occurred in the Midwest in 2009 (13). Our collection of field strains represents this recent emergence. Effector predictions in *Xg* field strains from this period revealed the presence of four potential effectors that are not present in the reference strain Xg 101 (Table 5). Xg field strains possess a XopJ1 that is identical to the allele in Xe85-10 and a Type III effector protein (T3EP) that has 78% amino acid identity to a predicted *Ralstonia* peptidase effector (WP_014619440.1). A predicted effector of the Xg strains shares 65% amino acid similarity to a *X. campestris* pv. *campestris* PTP Type III effector (WP_011345706.1). Two copies of XopE2 are present in 7 out of 10 Xg field strains, while the remaining three and the reference strain Xg 101 have only one XopE2. Two field strains carry the effector AvrBs7. Because the repetitive nature of TAL effector genes renders them difficult to assemble from short reads, Southern blot analysis was used to identify potential family members (Fig. 4-1A). In addition, the ability of each strain to induce a HR on pepper cv. ECW30R, which contains the cognate R gene *Bs3* to the TAL effector AvrHah1 was tested (Fig. 4-1B). All field strains of Xg contained a single TAL effector, an apparent AvrHah1, on the basis of band size and activity.

Although the Xg field strains were isolated within a three-year period, only three strains (Xg164, Xg165, and Xg167) have identical effector allele profiles at the nucleotide level (Fig. 2-3C). Three effectors are highly polymorphic: the *avrBs1*-class effector, of which three alleles were detected, and the two *xopE2* effectors, of which five and three alleles were detected. Two alleles of *xopAD* are present at equal frequencies in the Xg field strains, with both alleles present in field strains isolated in the same year in the same state (e.g. Xg165 and Xg173, Ohio 2011) and in the same year in different states (e.g. Xg153 and Xg156, Ohio and Michigan, respectively, 2010).

Common effectors between species

Effector predictions of the field strains has identified two new common putative effectors to add to the previously described list of eleven effectors shared between Xe, Xp, and Xg (6). XopE2 was identified in all Xp field strains and, while not in the reference Xp91-118, should, therefore, be considered a commonly shared core effector with Xe and Xg. The identification of AvrBsT in the majority of Xp field strains and an identical copy of Xe XopJ1 in Xg field strains indicates the presence of a more broadly defined YopJ-family effector to the commonly shared effector list.

Table 1. Summary of *Xanthomonas* field strains sequenced

Species	Strain name	Origin in US	Host isolated	Year isolated	Isolation ID	Collector
X. euvesicatoria	Xe073	North Carolina	P	1994	181	DR
	Xe074	Raleigh, NC	P	1994	199	DR
	Xe075	Soutwest FL	P	1995	206	DR
	Xe076	Naples, FL	P	1995	259	DR
	Xe077	Kentucky	P	1996	315	DR
	Xe078	Clewiston, FL	P	1997	329	DR
	Xe079	Jupiter, FL	P	1998	354	DR
	Xe081	Ft. Pierce, FL	P	1995	376	DR
	Xe082	Southeast FL	P	1998	455	DR
	Xe083	Belle Glade, FL	P	1999	490	DR
	Xe085	Boynton Beach, FL	P	1999	515	DR
	Xe086	Delray Beach, FL	P	2000	526	DR
	Xe091	Boca Raton, FL	P	2003	586	DR
	Xe101	Sampson Co., NC	P	2008	678	DR
	Xe102	Manetee, FL	P	2008	679	DR
	Xe103	Pender Co., NC	P	2009	681	DR
	Xe104	Sampson Co., NC	P	2010	683	DR
	Xe105	Granville, NC	P	2010	684	DR
	Xe106	Granville, NC	P	2010	685	DR
	Xe107	Granville, NC	P	2011	689	DR
	Xe108	Pender Co., NC	P	2012	695	DR
	Xe109	Cook Co., GA	P	2004	F4-2	DR
	Xe110	Tift Co. GA	P	2004	G4-1	DR
	Xe111	Colquitte Co., GA	P	2004	H3-2	DR
	Xe112	Brooks Co., GA	P	2004	L3-2	DR
X. perforans	Xp4B	Citra, FL	T	1998	Xp4B	JJ
	Xp2010	Hendry County, FL	P	2010	Xp2010	JJ
	TB6	Hillsborough, FL	T	2013	TB6	JJ
	TB9	Hillsborough, FL	T	2013	TB9	JJ
	TB15	Hillsborough, FL	T	2013	TB15	JJ
	Xp3-15	Decatur Co., GA	T	2006	Xp3-15	JJ
	Xp4-20	Decatur Co., GA	T	2006	Xp4-20	JJ
	Xp5-6	Decatur Co., GA	T	2006	Xp5-6	JJ
	Xp7-12	Manatee Co., FL	T	2006	Xp7-12	JJ
	Xp8-16	Manatee Co., FL	T	2006	Xp8-16	JJ
	Xp9-5	Manatee Co., FL	T	2006	Xp9-5	JJ
	Xp10-13	Manatee Co., FL	T	2006	Xp10-13	JJ
	Xp11-2	Palm Beach Co, FL	T	2006	Xp11-2	JJ
	Xp15-11	Miami-Dade Co., FL	T	2006	Xp15-11	JJ
	Xp17-12	Collier Co., FL	T	2006	Xp17-12	JJ
	Xp18-15	Collier Co., FL	T	2006	Xp18-15	JJ
	GEV839	Hardee Co., FL	T	2012	GEV839	JJ
	GEV872	Immokalee, FL	T	2012	GEV872	JJ
	GEV893	Collier Co.	T	2012	GEV893	JJ
	GEV904	Hillsborough, FL	T	2012	GEV904	JJ

	GEV909	Collier Co.	T	2012	GEV909	JJ
	GEV915	Hillsborough, FL	T	2012	GEV915	JJ
	GEV917	Hillsborough, FL	T	2012	GEV917	JJ
	GEV936	Lee, FL	T	2012	GEV936	JJ
	GEV940	GCREC, FL	T	2012	GEV940	JJ
	GEV968	Manatee Co., FL	T	2012	GEV968	JJ
	GEV993	Hendry Co., FL	T	2012	GEV993	JJ
	GEV1001	Quincy, FL	T	2012	GEV1001	JJ
	GEV1026	West Coast, FL	T	2012	GEV1026	JJ
	GEV1044	Collier Co., FL	T	2012	GEV1044	JJ
	GEV1054	Manatee Co., FL	T	2012	GEV1054	JJ
	GEV1063	Collier Co., FL	T	2012	GEV1063	JJ
X. gardneri	Xg153	Gibsonburg, OH	T	2010	SM194-10	SM
	Xg156	Blissfield, MI	T	2010	SM177-10	SM
	Xg157	Blissfield, MI	T	2010	SM182-10	SM
	Xg159	Blissfield, MI	T	2010	SM220-10	SM
	Xg160	Blissfield, MI	T	2010	SM234-10	SM
	Xg164	Ottawa, OH	T	2011	SM406-11	SM
	Xg165	Ottawa, OH	T	2011	SM413-11	SM
	Xg173	Unknown, OH	T	2011	SM605-11	SM
	Xg174	Wayne Co., OH	T	2012	SM775-12	SM
	Xg177	Sandusky Co., OH	T	2012	SM795-12	SM

Table 1. DR = David Ritchie, JJ = Jeffery Jones, SM = Sally Miller.

Table 2. Average statistics for *de novo* genome assemblies

Species	Strains	Contig L50	Contig size	Contig number	Genome coverage	Genome size (bp)
X. euvesicatoria	25	177763	42357	149	118	5327013
X. perforans	32	256872	71988	91	133	5231128
X. gardneri	10	42631	23650	244	57	5296542

Table 2. Average genome coverage was estimated by dividing the total average bases sequenced by the average predicted genome size (in bp). Contig number indicates the average number of contigs assembled. Contig L50 is a median statistic indicating the length of the contig where half of the total assembly is present in contigs equal or larger to the L50 length.

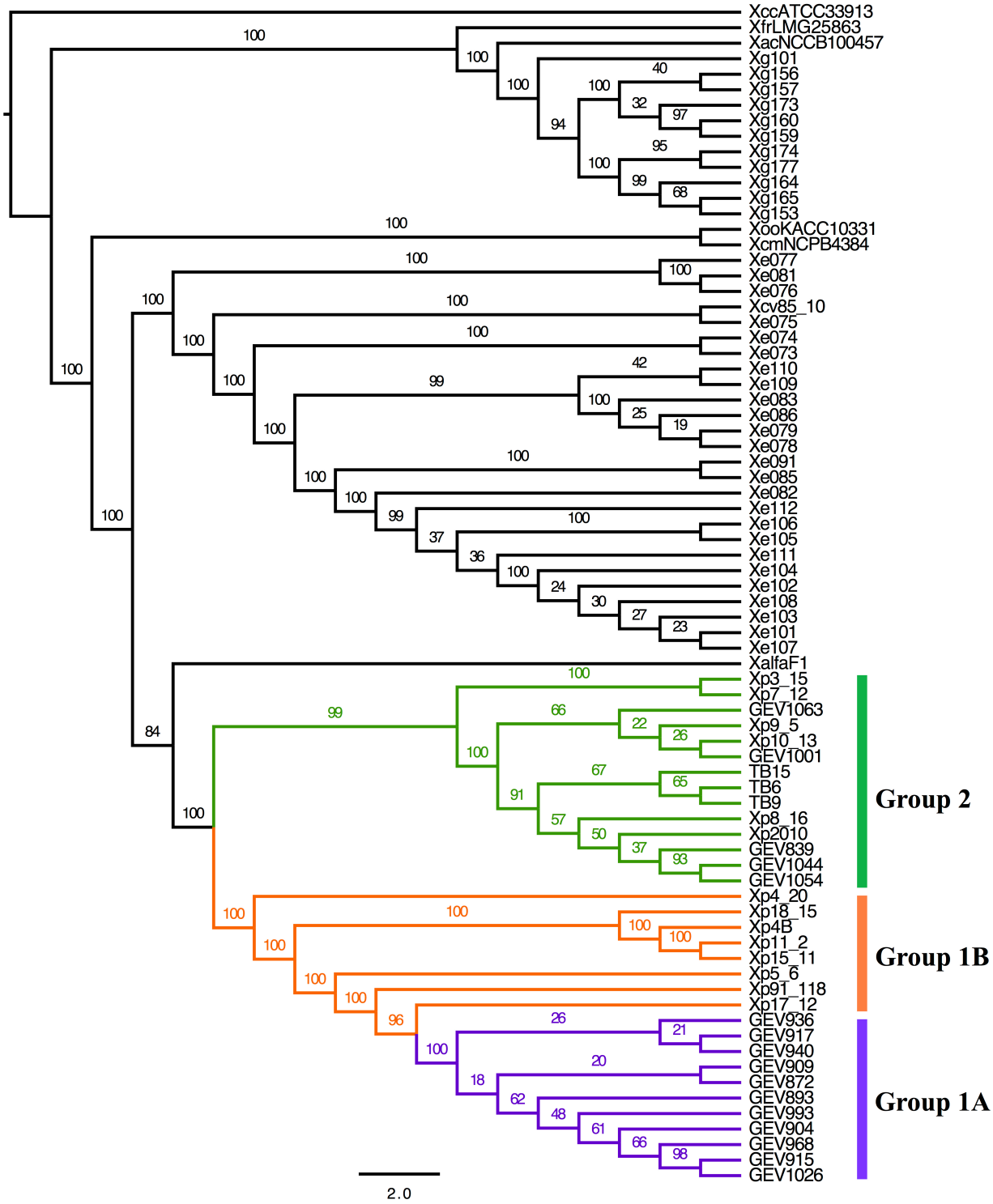


Figure 2-1. Core genome phylogenic analysis

Phylogenetic trees obtained by ML (IQTree) analysis, based on partitioned analysis (by codon position) of a total of 916,326 sites (1,017 genes families). Branch support values are shown for each tree, consisting of relative bootstrap proportions. Xp group designations are colored as follows: Group 1A: purple, Group 1B: orange, Group 2: green.

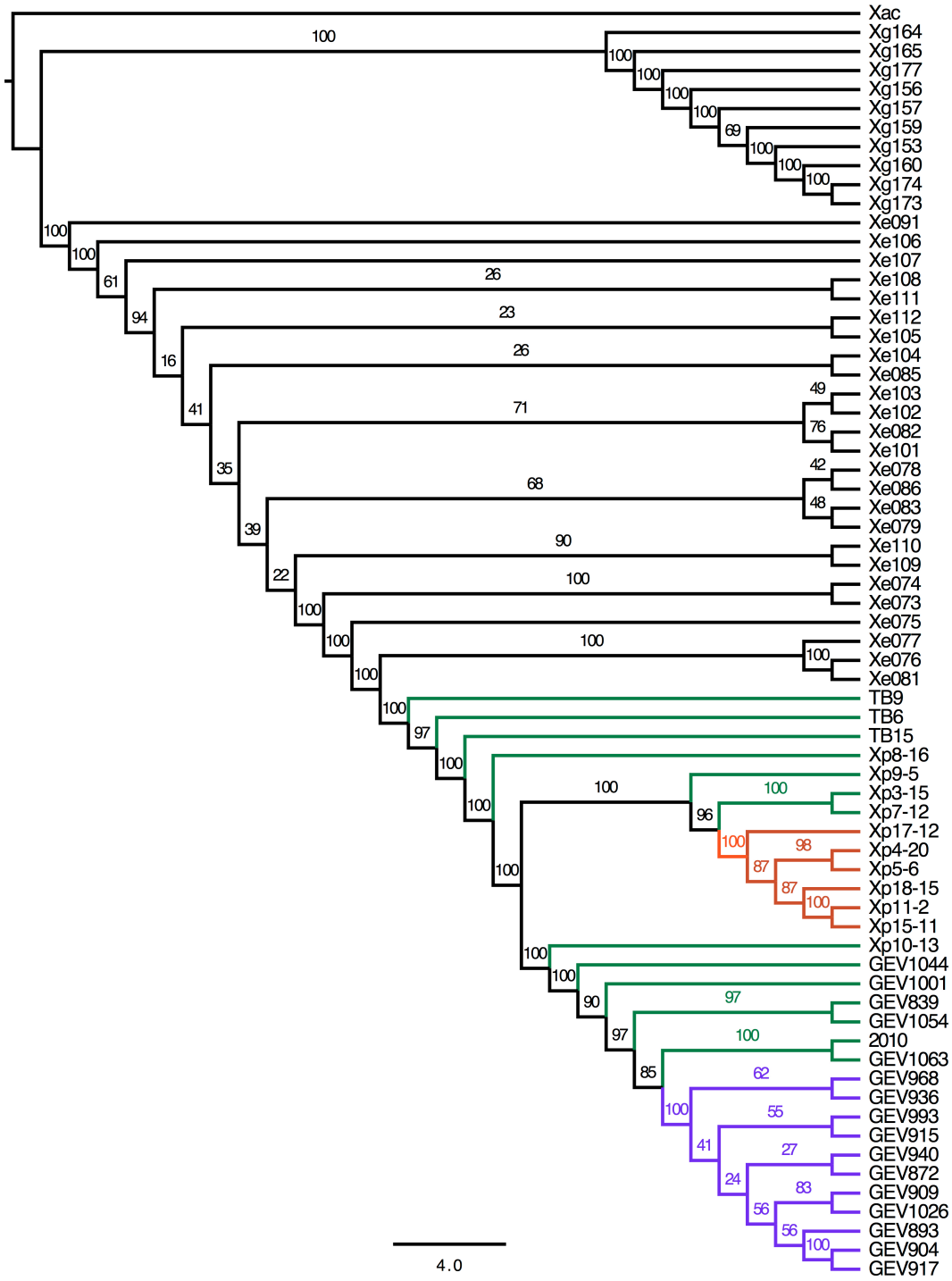


Figure 2-2. Phylogeny based on whole genome SNP analysis
Sequencing reads were mapped to *Xanthomonas axonopodis* pv. *citri* (Xac) reference number NC_003919 and bootstrap values are displayed. Scale bar corresponds to the number of nucleotide substitutions per site. Branches for Xp strains are colored to indicate group designations as in Figure 1: Group 1A: purple, Group 1B: orange, Group 2: green.

Table 3. *Xanthomonas euvesicatoria* nucleotide type III effector alleles

	year	1985	1994		1995			1996	1997	1998		1999		2000	2003	2004				2008		2009	2010			2011	2012
	state		NC	NC	F	F	F	K	F	F	F	F	F	F	F	G	G	G	G	NC	F	NC	NC	NC	NC	NC	NC
Type III Effector	strain	85-10	Xe073	Xe074	Xe075	Xe076	Xe081	Xe077	Xe078	Xe079	Xe082	Xe083	Xe085	Xe086	Xe091	Xe109	Xe110	Xe111	Xe112	Xe101	Xe102	Xe103	Xe104	Xe105	Xe106	Xe107	Xe108
		4	7	7	7	7	4	7	3	1	2	6	5	6	5	1	1	2	2	2	2	2	2	2	2	2	2
XopA		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
XopAA		1	2	2	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
XopAD		1	2	2	1	3	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
XopAE		0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
XopAF-like		0	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
XopAI		1	2	2	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
XopAK		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
XopB		1	1	1	1	2	2	2	1	1	1	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
XopC1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
XopD		1	1	1	1	2	2	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
XopE1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
XopE2		1	4	3	1	1	1	3	1	1	2	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2
XopE3		0	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
XopF1		1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
XopF2		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
XopG		1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
XopI		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
XopJ1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
XopJ3		1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1
XopK		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
XopL		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
XopN		1	2	2	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
XopO		1	2	2	2^{CTG}	1^{T}	1^{T}	1^{T}	2	2	2	2	2	2	2	2	3	2	2	2	2	2	2	2	2	2	2
XopP		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
XopQ		1	2	2	1	2	2	2	2	2	2	2	2	2	3	2	2	2	2	2	2	2	2	2	2	2	2
XopR		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
XopV		1	2	2	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
XopX		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
XopZ1		1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	3	1	1	1
PTP		0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 4. *Xanthomonas perforans* amino acid type III effector alleles

	year	1991	1998	2006											2010	2012																2013			1985
	state	F	F	G	G	G	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F
Type III effector	strain	Xp91-118	Xp4B	Xp3-15	Xp4-20	Xp5-6	Xp7-12	Xp8-16	Xp9-5	Xp10-13	Xp11-2	Xp15-11	Xp17-12*	Xp18-15	Xp2010*	GEV839	GEV872	GEV893	GEV904	GEV909	GEV915	GEV917	GEV936	GEV940	GEV968	GEV993	GEV1001	GEV1026	GEV1044	GEV1054	GEV1063	TB6	TB9*	TB15*	Xe 85-10
AvrBs2		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2
XopA		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2
XopAP		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2
XopAR		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
XopC2		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2
XopE1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	3	1	2	
XopF1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	
XopF2		1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	
XopI		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	
XopJ4		1	1	2	1	1	2	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	
XopK		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	
XopL		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	
XopN		1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	3	
XopQ		1	1	2	1	1	2	2	2	2	1	1	1	1	2	2	1	1	1	1	1	1	1	1	1	1	1	2	1	2	2	2	2	2	2
XopR		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	
XopV		1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	3	1	1	1	1	1	1	1	1	4	
XopX		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	
XopAK		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	
XopD		3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	4	5	4	4	4	4	4	4	4	4	7	4	3	6	3	3	7	3	8
XopAD		1	1	1	1	1	1	1	1	1	1	1	3	3	1	1	1	1	1	1	1	1	1	2	1	1	1	2	1	1	1	1	1	1	3
XopE2		0	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1
XopE4		0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AvrBsT		0	1	1	1	0	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
XopP1		1	1	1	1	0	3	4	0	0	5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
XopP2		2	5	17	2	4	8	6	9	0	3	1	1	18	15	14	1	1	1	1	1	1	1	1	1	1	1	10	1	11	12	13	19	19	19	16
XopP3		0	0	0	0	0	1	1	1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
XopZ1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2
XopAE		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2
XopAF		1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Group		1B	1B	2	1B	1B	2	2	2	2	1B	1B	1B	1B	2	2	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	1A	2	1A	2	2	2	2	2	2	

Table 5. *Xanthomonas gardneri* nucleotide type III effector

	year	1953	2010					2011			2012	
	state		O	M	M	M	M	O	O	O	O	O
Type III Effector	Strain	Xg101	Xg153	Xg156	Xg157	Xg159	Xg160	Xg164	Xg165	Xg173	Xg174	Xg177
AvrBs1 class		1	2	3	3	3	3	2	2	3	2	2
AvrBs2		1	1	1	1	1	1	1	1	1	1	1
AvrBs7		0	0	1	0	0	0	0	0	1	0	0
AvrHah1		na	na	na	na	na	na	na	na	na	na	na
AvrXccA1		1	1	1	1	1	1	1	1	1	1	1
XopAD		1	1	2	2	2	2	1	1	2	1	1
XopAM		1	1	1	1	1	1	1	1	1	1	1
XopAO		1	1	1	1	1	1	1	1	1	1	1
XopAQ		1	1	0	0	1	1	1	1	1	1	1
XopAS		1	1	1	1	1	1	1	1	1	1	1
XopB		1	1	1	1	1	1	1	1	1	1	1
XopD		1	1	1	1	1	1	1	1	1	1	1
XopE2_0		1	1	4	5	3	1	5	5	4	2	5
XopE2_1		0	1	1	3	2	1	0	0	1	1	0
XopF1		1	1	1	1	1	1	1	1	1	1	1
XopG		1	FS	1	1	1	1	1	1	1	1	1
XopJ1		0	1	1	1	1	1	1	1	1	1	1
XopK		1	1	1	1	1	1	1	1	1	1	1
XopL		1	1	1	1	1	1	1	1	1	1	1
XopN		1	1	1	1	1	1	1	1	1	1	1
XopQ		1	1	1	1	1	1	1	1	1	1	1
XopR		1	1	1	1	1	1	1	1	1	1	1
XopX		1	1	1	1	1	1	1	1	1	1	1
ZopZ2		1	1	1	2	1	1	1	1^{CTG}	1	1	1
PTP		0	1	1	1	1	1	1	1	1	1	1
T3EP		0	1	1	1	1	1	1	1	1	1	1

Table 3. ***Xanthomonas euvesicatoria*** **nucleotide type III effector alleles** Each distinct nucleotide allele was assigned an arbitrary number. The number 0 indicates the effector is missing from genomic assemblies. Superscripts are as follows: T = truncation, I = insertion + base pair number, CTG = contig break in assembly unable to be confirmed via Sanger Sequencing. PTP = protein tyrosine phosphatase.

Table 4. ***Xanthomonas perforans*** **amino acid type III effector alleles.** Each distinct amino acid allele was assigned an arbitrary number. The number 0 indicates the effector is missing from genomic assemblies. AvrBsT and XopP were not used to make the Neighbor-joining effector allele trees for Xp in Figure 3A. Asterisk (*) indicates newly identified Xp pepper pathogens.

Table 5: ***Xanthomonas gardneri*** **nucleotide type III effector alleles.** Each distinct nucleotide allele was assigned an arbitrary number. The number 0 indicates the effector is missing from genomic assemblies CTG = contig break in assembly unable to be confirmed via Sanger Sequencing. FS indicates a frame shift mutation. The TAL effector AvrHah1 could not be assembled (na = not assembled). PTP = protein tyrosine phosphatase, T3EP = type III effector protein.

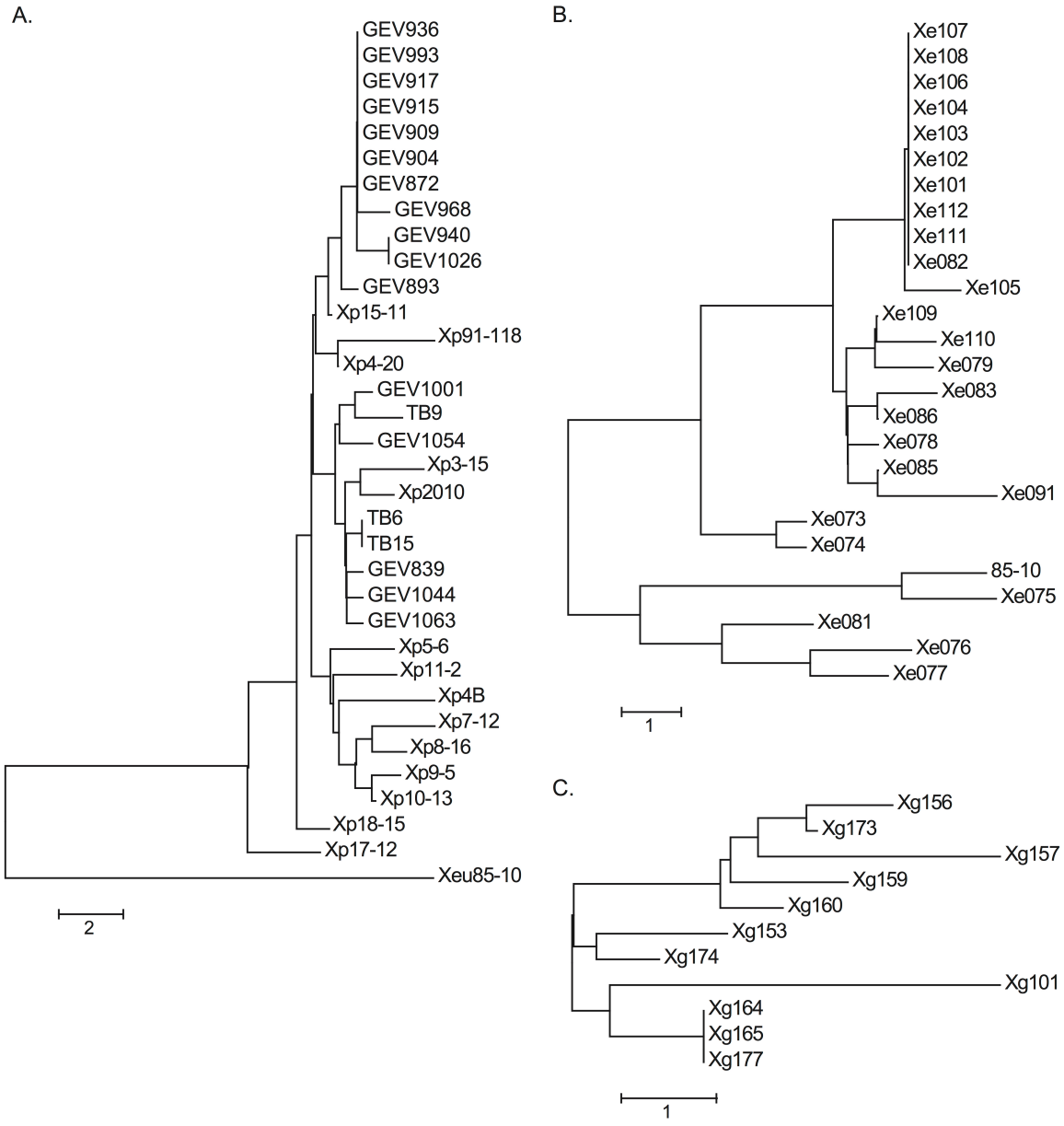


Figure 2-3. Neighbor-joining trees of effector allele profiles
Neighbor-joining trees for (A) Xp (B) Xe and (C) Xg field strains were constructed using nucleotide (Xe and Xg) and amino acid (Xp) pairwise allele differences between strains. Effector allele designations can be found in Tables 3-5. A difference at one effector between two strains equals a distance of 1.0.

Discussion

The population dynamics of *Xanthomonas*-infecting pepper and tomato has shifted in the United States over the past twenty-five years. Prior to 1991, Xe was the prevalent species and the only species in tomato fields in Florida. Xp tomato race 3 was identified first in 1991 and eventually replaced the Xe population in tomato fields, a process attributed to the ability of Xp race 3 to out-compete Xe via the production of bacteriocins (10, 11). Xp4B, a tomato race 4 strain identified in 1998, carries a mutation in the *avrXv3* gene. Field surveys thereafter in 2006 and 2012 recovered a majority of race 4 strains carrying either frameshift mutations or transposon insertions in *avrXv3* (12). The first reports of Xg in the United States occurred in Ohio and Michigan during a bacterial spot outbreak on tomato in 2009 (13).

We sequenced Xe, Xp, and Xg strains isolated in different years, from different fields/transplant houses throughout southeastern and midwestern United States. We have also sequenced strains collected during the same season from the same field. Following typical population genomic studies, we have taken three components into consideration; location, time and niche (43). Combining genomic data with metadata such as plant host source, year and location of isolation provides inference of population structure and clues to host adaptation. We have computationally predicted the Type III effector repertories for each strain, and have used two different methods in order to infer evolutionary relationships of strains based on whole genome data. Phylogeny based on the core genome considers orthologous genes among the set of genomes considered. Phylogeny based on whole genome SNPs included core as well as variable regions of the genome, and thus provides an additional method to describe the genetic diversity within field strains. Phenotypic data, in particular, host range, was then correlated with the whole genome phylogenies.

MLSA studies showed the presence of two distinct groups of Xp populations that appeared to be clonal within the lineage (8). However, these studies were based on 6 genes out of 5000 genes. In our study, core ortholog gene phylogeny also revealed two distinct groups among Xp populations (Groups 1 and 2), although we were able to further separate Group 1 into Group 1A and 1B. Group 1A contains eleven strains isolated in 2012, whereas Group 1B contains six strains isolated in 2006 in addition to Xp4B and the reference strain Xp91-118, isolated in 1998 and 1991, respectively. Group 2 comprises five strains from 2006, the single strain from 2010, five strains from 2012, and all three 2013 strains. Additionally, we detected genetic diversity among strains that appeared to be clonal from MLSA in previous work (8), particularly evident in Xp Group 1A. In our study, the Xp 2006 population was more diverse than the 2012 population, possibly due to the fact that sampling in 2006 was carried out in a broader geographic range in Florida and Georgia. The diversity within the 2006 population is evident from the core genome and SNP phylogenies.

This work re-emphasizes the role of population genomics for identification of elements involved in host-pathogen arms race. The data revealed the emergence of tomato race 4 strains of Xp carrying mutations (either frameshift/transposon insertion) in *avrXv3.* Strain Xp91-118 isolated in 1991 was non-pathogenic on pepper even when mutated in

avrXv3 (44), indicating the existence of other factors that restrict its host range on pepper. We will explore additional effectors that restrict host range on pepper (and *Nicotiana benthamiana*) in Section 3.

Field strain genomic analysis with computational Type III effector predictions are an efficient method for deriving the diversity of Type III effector repertories (and discovery of newly acquired effectors) compared to traditional PCR approaches. Knowledge of the effector content in the population will inform strategies for achieving broad durable resistance strategies based on R gene deployment. Within each species, we identified several differences in the effector repertoires of Xe, Xp, and Xg field strains, including the gain or loss of effector genes, null mutations, and the presence of alternate alleles. We predicted three effector additions to the overall Xe field strain repertoire (XopE3, XopAF-like, and XopAE) and one removal (XopG) in comparison to the reference strain Xe85-10. Unsurprisingly, the most polymorphic effector in Xe is *avrBs2*, a phenomenon perhaps explained by the selective pressures of the pepper *Bs2* resistance gene deployed in the early 1990's. Several of the previously reported mutations in *avrBs2* are represented here, with no novel polymorphisms detected (45, 46). Generally, the effector predictions for Xe field strains isolated between 1994 and 2004 show increased effector polymorphisms compared to strains isolated between 2004-2012, indicating that the effector repertoires have stabilized over time in our sampling population. Xp field strains have evolved their repertoires by losing/gaining effectors (XopE2, XopE4, AvrBsT), through allelic exchange (as seen with XopQ in Group 2 strains) and by frameshift mutations/transposon insertions (in *avrXv3*). Diversity in effector repertoires is seen even in strains collected from the same field during a single growing season. Strains TB6 and TB15 possess identical Type III effector profiles and appear clonal based on core genome phylogeny except for the absence of AvrBsT in TB15. Similar to TB15, TB9 does not posses AvrBsT but has different alleles of XopD and XopE1 compared to TB6 and TB15. We predicted four additions to the Xg field strain effector repertoire including a second copy of XopE2 and a XopJ1 identical to Xe strains. We also detected allele differences in an AvrBs1-like effector, XopAD, and XopE2. Through this analysis two additional effectors can be added to the previous list of eleven commonly shared effectors between Xe, Xp, and Xg (6): XopE2 and a YopJ-family member (AvrBsT in Xp, XopJ1 in Xg and Xe).

In addition to strain-level variation, allelic diversification in Type III effectors was observed at the species level across Xe, Xp, and Xg. Because Type III effector repertoires are proposed to be a major factor determining host range (22), it is important to understand the diversity of effectors present in different species that infect common hosts. Although Xe, Xp, and Xg share thirteen core effectors, effector alleles between these three species may be considerably different. For example, the effector AvrBs2 protein sequence shares 99% identity between reference strains Xp91-118 and Xe85-10, but 77% identity to the AvrBs2 in Xg101. Similarly, the XopQ alleles of Xp91-118 and Xe85-10 share 99% identity at the amino acid level, but 58% identity to XopQ from Xg 101. Sampling of a genetically diverse population can be informative to reveal the dominant effector alleles in a specific geographical region, which would be the most appropriate alleles to screen for R protein resistance strategies.

The increased speed and reduction of cost of DNA sequencing technology combined with the use of genome editing techniques are providing new opportunities for designing resistance strategies against specific pathogens in various crop species. The spread of

agricultural pathogens into new niches, either by increasing global movement of crop plants or the emergence of new niches from climate change, makes the continued genomic surveillance of agricultural pathogens a top priority for food security and resistance strategies. Of particular importance are tracking shifts in dominant species and changes in effector repertoires and alleles. Effector maintenance and stability is a key consideration for the future design of durable resistance strategies using R-gene employment into crops.

3. Two Type III effectors define host range for Xe and Xp, but not Xg

Background

Xp has previously been considered restricted to tomato as a host. In 2010, a Xp strain from a greenhouse-grown diseased pepper plant was isolated and confirmed as Xp based on 16S rRNA sequencing and multilocus sequence analysis (MLSA) (8). This strain, designated here as Xp2010, does not induce a hypersensitive response (HR) on pepper cv. Early CalWonder (ECW) and is able to create foliar disease lesions. Effector predictions for Xp2010 from section 2 of this work indicated that the absence of AvrBsT, which induces HR on pepper (47), may be responsible for its pepper host expansion. We were curious to see if other Xp strains in our collection displayed host expansions to pepper similar to Xp2010 and if this could be explained solely by the presence of AvrBsT. Our analysis shows that this is true for Xp strains in Group 2, but not for Group 1A and 1B strains, which gain only partial virulence on pepper when mutated in *avrBsT.*

The ability of a single effector to tip the balance determining host and non-host interactions with plants is also true in the case of in hopQ1-1 in *Pseudomonas syringae* pv. *tomato* (Pto) DC3000 for *Nicotiana benthamiana* (48). Normally a pathogen of tomato and *Arabidopsis thaliana*, a DC3000ΔhopQ1-1 mutant is able to cause bacterial speck symptoms on *N. benthamiana*. Because the *Xanthomonas* effector XopQ is a homolog of Pto HopQ1-1, we hypothesized that deletion of XopQ in Xe, and a double deletion of XopQ and AvrBsT in Xp, would expand the host ranges of Xe and Xp to cause disease in *N. benthamiana.* Additionally, we discovered that Xg, which does not contain AvrBsT and has a unique allele of XopQ from Xe and Xp is a natural pathogen of *N. benthamiana*.

Results

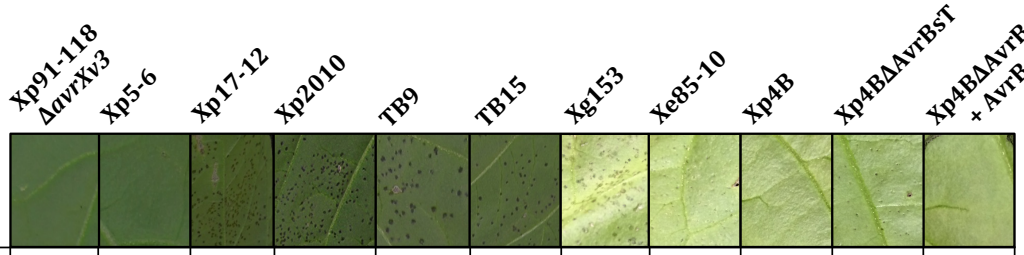
Association of AvrBsT presence or absence in host range expansion of *Xp* on pepper

Because we observed that Xp2010, which lacks the effector AvrBsT, is a pepper pathogen, we wondered if other Xp strains would show the same gain of host expansion to pepper due to the loss of AvrBsT. We first made an AvrBsT deletion in Xp4B, which is used as an experimental type strain in our laboratory. Normally non-pathogenic on pepper, a deletion mutant of AvrBsT in Xp4B caused weak disease lesions on pepper and gained partial virulence—measured by *in planta* growth—compared to pepper pathogens Xg153 and Xe85-10 (Fig. 3-1, Fig. 3-2). For comparison, the growth level of XgΔhrcV, which does not secrete any Type III effectors and induces basal disease resistance, or PAMP Triggered Immunity (PTI), is shown (Fig. 3-2). Complementation of AvrBsT back into Xp4BΔAvrBsT prevented lesion formation and reduced *in planta* growth back to restricted wt levels.

We used PCR to confirm the presence or absence of AvrBsT in the *Xp* field strains, using Xp4B and Xp2010 as positive and negative controls, respectively (Table 6). Inoculation of Xp4B into pepper at a high inoculum (10^8 CFU/ml) induces the plant immune response, or HR, while the pathogenic Xp2010 induces water soaking, a sign of disease. We similarly infiltrated the Xp field strains into ECW and recorded whether HR or water soaking occurred (Table 6). We confirmed that four additional field strains, Xp5-6, Xp17-12, TB9 and TB15 do not possess AvrBsT and also fail to induce HR on ECW. Of these strains, Xp17-12, TB9, and TB15, but not Xp5-6, are able to cause disease lesions on pepper cv. ECW when infiltrated at a low inoculum (10^4 CFU/ml). Strain Xp5-6 showed a phenotype similar to Xp91-118Δ*avrXv3*, which is unable to cause lesions on pepper, indicating that additional factors restrict the host range of Xp5-6 on pepper (Fig. 3-1).

Xp AvrBsT mutants in Group 2, but not Group 1A or 1B, gain full pepper virulence

Three of the newly identified Xp pepper pathogens that lack AvrBsT—Xp2010, TB9, and TB15—belong to Group 2. Xp4B, a Group 1B strain, gained only partial virulence on pepper when mutated in *avrBsT*. Similarly, Xp5-6, a Group 1B strain, lacks AvrBsT but does not cause disease on pepper. We hypothesized that Group 2 strains carrying mutations in AvrBsT would exhibit *in planta* growth and virulence similar to that of virulent strains from pepper in our collection. At the same time, strains belonging to Group 1 and carrying mutations in *avrBsT* would be non-pathogenic on pepper, similar to Xp91-118Δ*avrXv3* (44).



	Xp91-118 ΔavrXv3	Xp5-6	Xp17-12	Xp2010	TB9	TB15	Xg153	Xe85-10	Xp4B	Xp4BΔAvrBsT	Xp4BΔAvrBsT + AvrBsT
Disease	-	-	+	+	+	+	+	+	-	+	-
AvrBsT	-	-	-	-	-	-	-	-	+	-	+
Group	1B	1B	1B	2	2	2	n/a	n/a	1B	1B	1B

Figure 3-1. Xp disease phenotypes on pepper ECW
Xp strains were syringe infiltrated at 10^4 CFU/mL and lesions were allowed to develop for 8-10 days. Presence or absense of disease lesions are indicated with a + and -, respectively. Xp group designations were determined by core genome phylogeny (Fig. 2-1).

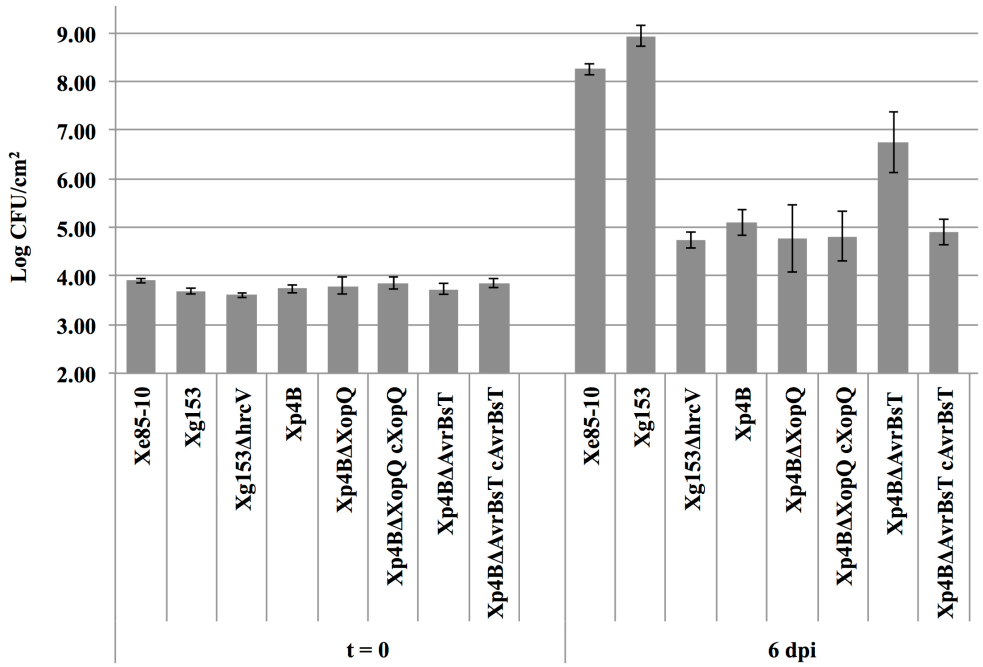


Figure 3-2. Xp4BΔAvrBsT does not gain full virulence on pepper cv. ECW
In planta growth of the strains above were measured at 6 days post infiltration from a starting concentration of 10^5 CFU/mL.

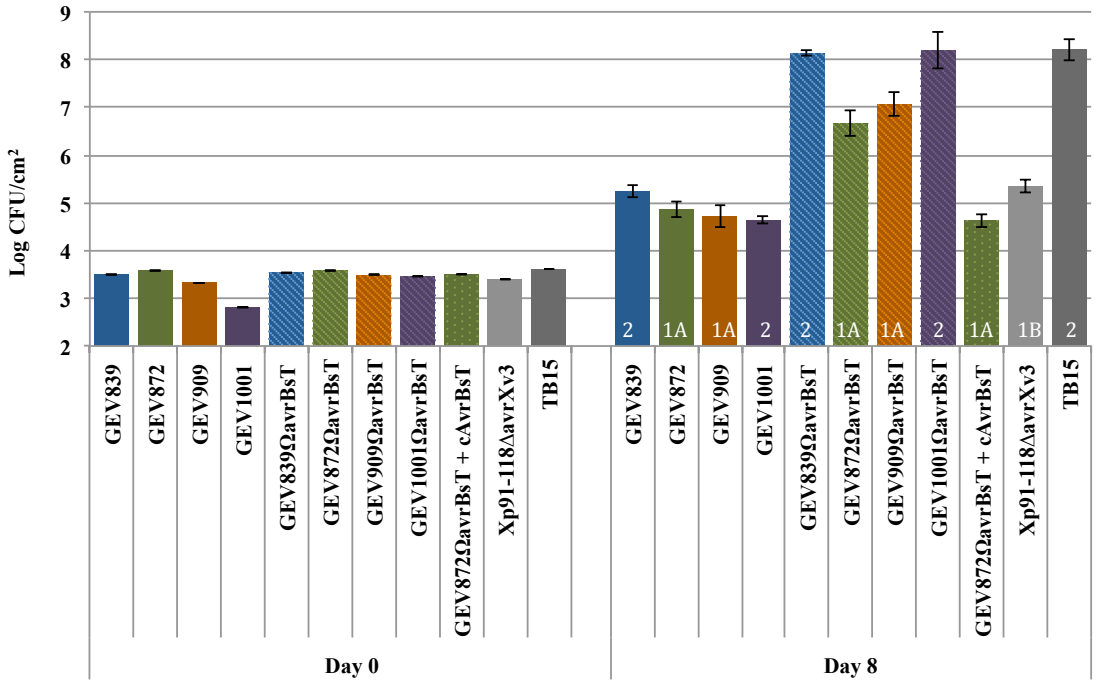


Figure 3-3. Group 2, but not Group 1, AvrBsT mutants gain full virulence on pepper
In planta growth of *X. perforans* strains and *avrBsT* insertion mutants. Group designations are marked in white over Day 8 growth.

To test this hypothesis, *avrBsT* insertion mutants (designated as ΩAvrBsT) were introduced into two Group 2 strains, GEV839 and GEV1001, and two Group 1A strains, GEV872 and GEV909 and tested for *in planta* growth (Fig. 3-3) and high inoculum HR or water soaking (Table 6) on pepper cv. ECW.

Indeed, XpGEV839Ω*avrBsT* and XpGEV1001Ω*avrBsT* from Group 2 lose the ability to elicit HR in pepper and are virulent similar to the natural Group 2 pepper pathogen TB15 (Fig. 3-3, Table 6). *In planta* population levels for these two mutants were not significantly different from TB15 at Day 8 post-infiltration, indicating that AvrBsT is the lone factor restricting these two Group 2 strains on pepper. Also as predicted, insertion mutants of *avrBsT* in Group 1A strains GEV872 and GEV909 lose the ability to induce HR on pepper but do not grow to the same extent as TB15. *In planta* populations of XpGEV872Ω*avrBsT* and XpGEV909Ω*avrBsT* were 100 fold higher compared to 91-118Δ*avrXv3* but 20-50 fold lower compared to pepper pathogens XpGEV839Ω*avrBsT*, XpGEV1001Ω*avrBsT* and TB15, indicating the existence of additional factors restricting the virulence of Group 1A strains on pepper. Complementation of XpGEV872Ω*avrBsT* with AvrBsT restricted the *in planta* growth to XpGEV872 levels.

Table 6: AvrBsT in *Xanthomonas perforans* and its role in host specificity on pepper (next page). Xp strains are organized according to isolation year. The presence of AvrBsT was first determined with effector predictions and then confirmed with PCR and HR on pepper cv. ECW. The pathogenicity of Xp strains on pepper is described after infiltration at a low inoculum to observe lesion development. The Group designations for Xp strains are described according to whole genome ortholog analysis. NT: not tested, n/a: not applicable. The virulence of *avrBsT* mutants was tested on pepper cv. ECW using growth assays or high inoculum (10^8 CFU/mL) spots to show water soaking. Asterisk (*) indicates newly identified Xp pepper pathogens.

Table 6. AvrBsT in *Xanthomonas perforans* and its role in host specificity on pepper

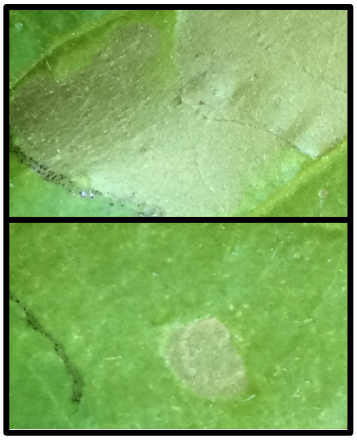
Isolated	Strain	HR on ECW	Disease on ECW	AvrBsT	Group	ΔAvrBsT mutant phenotype on ECW
1991	Xp91-118	-	No leasions	-	1B	NT
1998	Xp4B	+	No lesions	+	1B	not fully virulent *in planta*, no 10^8 infiltration-watersoaking phenotype, weak lesions
2006	Xp3-15	+	NT	+	2	NT
	Xp4-20	+	NT	+	1B	NT
	Xp5-6	-	weak lesions, similar to Xp91-118 ΔavrXv3	-	1B	natural AvrBsT mutant
	Xp7-12	+	NT	+	2	NT
	Xp8-16	+	NT	+	2	NT
	Xp9-5	+	NT	+	2	NT
	Xp10-13	+	NT	+	2	NT
	Xp11-2	+	NT	+	1B	NT
	Xp15-11	+	NT	+	1B	NT
	Xp17-12*	-	Disease lesions	-	1B	natural AvrBsT mutant
	Xp18-15	+	NT	+	1B	NT
2010	Xp2010*	-	Disease lesions	-	2	10^8 infiltration-watersoaking phenotype
2012	GEV839	+	No lesions	+	2	10^8 infiltration-watersoaking phenotype
	GEV872	+	No lesions	+	1A	*In planta* growth 100 fold higher compared to 91-118 avrXv3 mutant and 50 fold lower compared to pepper pathogen, TB15.
	GEV893	+	No lesions	+	1A	NT
	GEV904	+	No lesions	+	1A	NT
	GEV909	+	No lesions	+	1A	*In planta* growth 100 fold higher compared to 91-118 avrXv3 mutant and 10-20 fold lower compared to pepper pathogen, TB15
	GEV915	+	No lesions	+	1A	NT
	GEV917	+	No lesions	+	1A	NT
	GEV968	+	No lesions	+	1A	NT
	GEV993	+	No lesions	+	1A	NT
	GEV1001	+	No lesions	+	2	10^8 infiltration-watersoaking phenotype, *in planta* growth similar to pepper pathogen TB15.
	GEV1026	+	No lesions	+	1A	NT
	GEV1044	+	No lesions	+	2	NT
2013	TB6	+	No lesions	+	2	NT
	TB9*	-	Disease lesions	-	2	natural AvrBsT mutant
	TB15*	-	Disease lesions	-	2	natural AvrBsT mutant

Loss of XopQ and AvrBsT expands the host range of *Xp* to *Nicotiana benthamiana*

Members of both the XopQ and AvrBsT effector families are known to induce a HR in *N. benthamiana* (47, 48), indicating the existence of R-proteins in *N. benthamiana* that respond to these two effectors. Previous work had described the host gain of Pto DC3000, normally a pathogen of tomato and *Arabidopsis thaliana*, to *N. benthamiana* due to the deletion of a single effector, the XopQ homolog hopQ1-1. We hypothesized that a gene deletion of *xopQ* in Xe would similarly expand the host range of Xe to include *N. benthamiana*. We further hypothesized that a double deletion of *xopQ* and *avBsT* would expand the host range of Xp to include *N. benthamiana*. We created markerless whole gene deletions and tested mutant strains for pathogenicity on *N. benthamiana* using *in planta* growth assays, low inoculum single lesion assays, and high inoculum HR or water soaking assays (Fig. 3-4).

As expected, wt Xe85-10 is not pathogenic on *N. benthamiana* and causes a weak HR. However, a deletion of *xopQ* (Xe85-10ΔXopQ) results in water soaking, disease lesions, and high *in planta* growth after 6 days. Complementation of Xe85-10ΔXopQ with XopQ (on plasmid pVSP61 with the native promoter) restored the original avirulent phenotype of low *in planta* growth, no lesions, and HR, indicating that, like hopQ1-1 in PtoDC3000, XopQ is the sole avirulence factor restricting Xe pathogenicity on *N. benthamiana*.

To test the effect of a XopQ and AvrBsT double deletion on Xp host expansion, we created single and double mutants in an Xp4B background. Single gene knockouts for *xopQ* and *avrBsT* in Xp4B remained incompatible on *N. benthamiana* (HR, low growth, no lesions). However, the double effector deletion Xp4BΔXopQΔAvrBsT gave disease lesions at a low inoculum, showed water soaking at a high inoculum, and grew to high levels (Fig. 3-4). These results confirm our hypothesis and indicate that mutants in *xopQ* and *avrBsT* are sufficient to expand the host range of Xp to *N. benthamiana*.



Curiously, we discovered that Xg is a natural pathogen of *N. benthamiana*. Although XopQ is a core effector in Xe, Xp, and Xg, the Xg XopQ protein shares only 61% amino acid identity with XopQ from Xe and Xp. (Fig. 3-5). Due to this high degree of amino acid sequence variation, Xg XopQ may be unrecognized by the potential R-protein in *N. benthamiana* responsible for recognition of Xe and Xp XopQ. Because Xe induces a weak HR in *N. benthamiana*, we tested the XopQ alleles in *Nicotiana tabacum.* We observed strong HR in response to Xe XopQ, but no HR response to Xg XopQ (Fig. 3-6).

Figure 3-6. The XopQ allele from Xe, but not Xg, induces HR in *Nicotiana tabacum*
Transient expression from *Agrobacterium* carrying XopQ from Xe or Xg 48 hpi, OD_{600} = 0.1

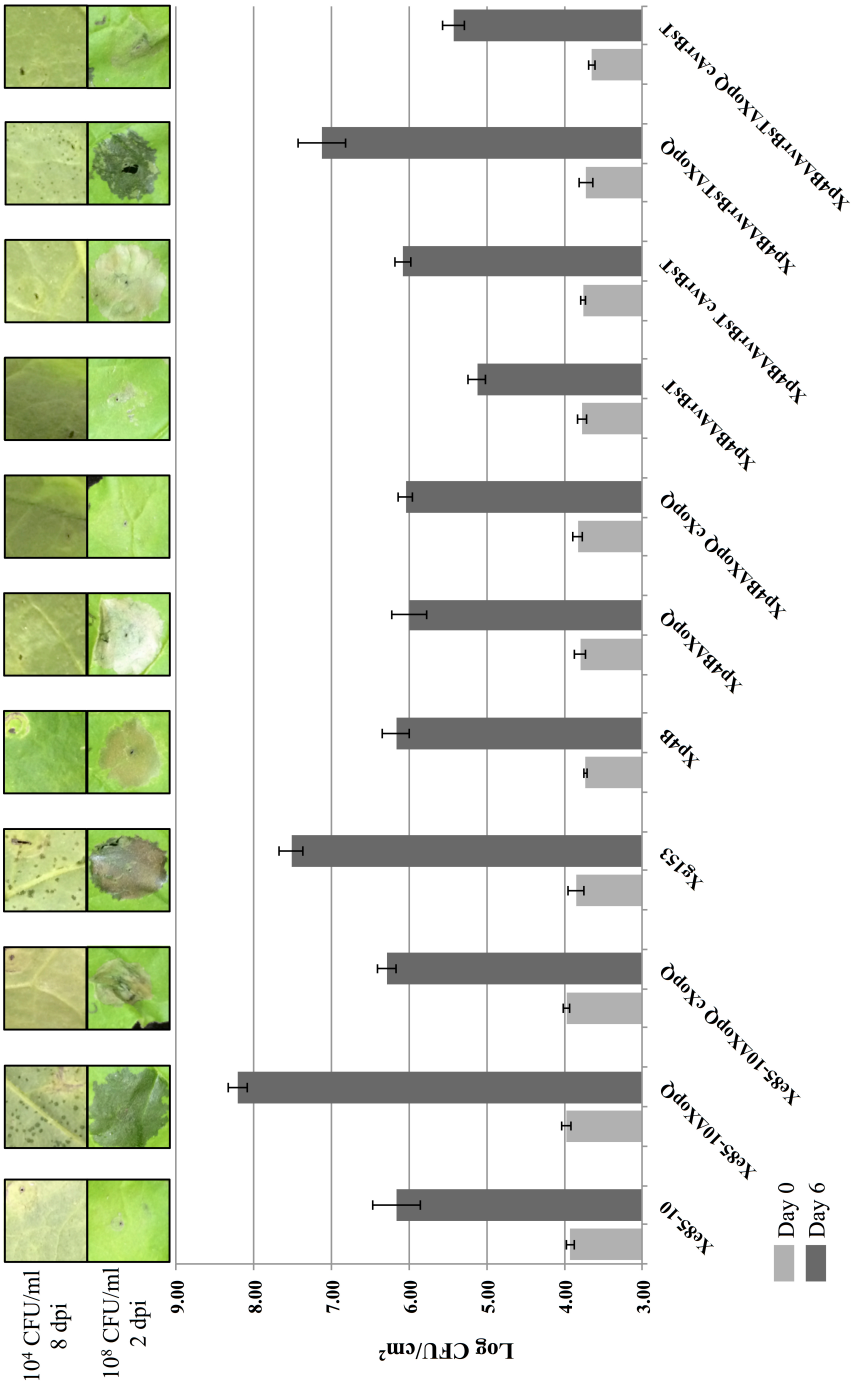


Figure 3-4. Host expansion of *Xanthomonas* spp. on *Nicotiana benthamiana*
In planta growth of *avrBsT* and *xopQ* deletion mutants measured at Day 0 and Day 6 with a starting inoculum of 10^5 CFU/mL after infiltration of leaves of *Nicotiana benthamiana*. Infiltrations on *N. benthamiana* were performed at 10^4 CFU/ml to display lesions and photographed 8 days post infiltration (8 dpi). High inoculum infiltrated spots were performed at 10^8 CFU/ml to show HR or water soaking and photographed 2 dpi. This experiment was repeated three times with similar results.

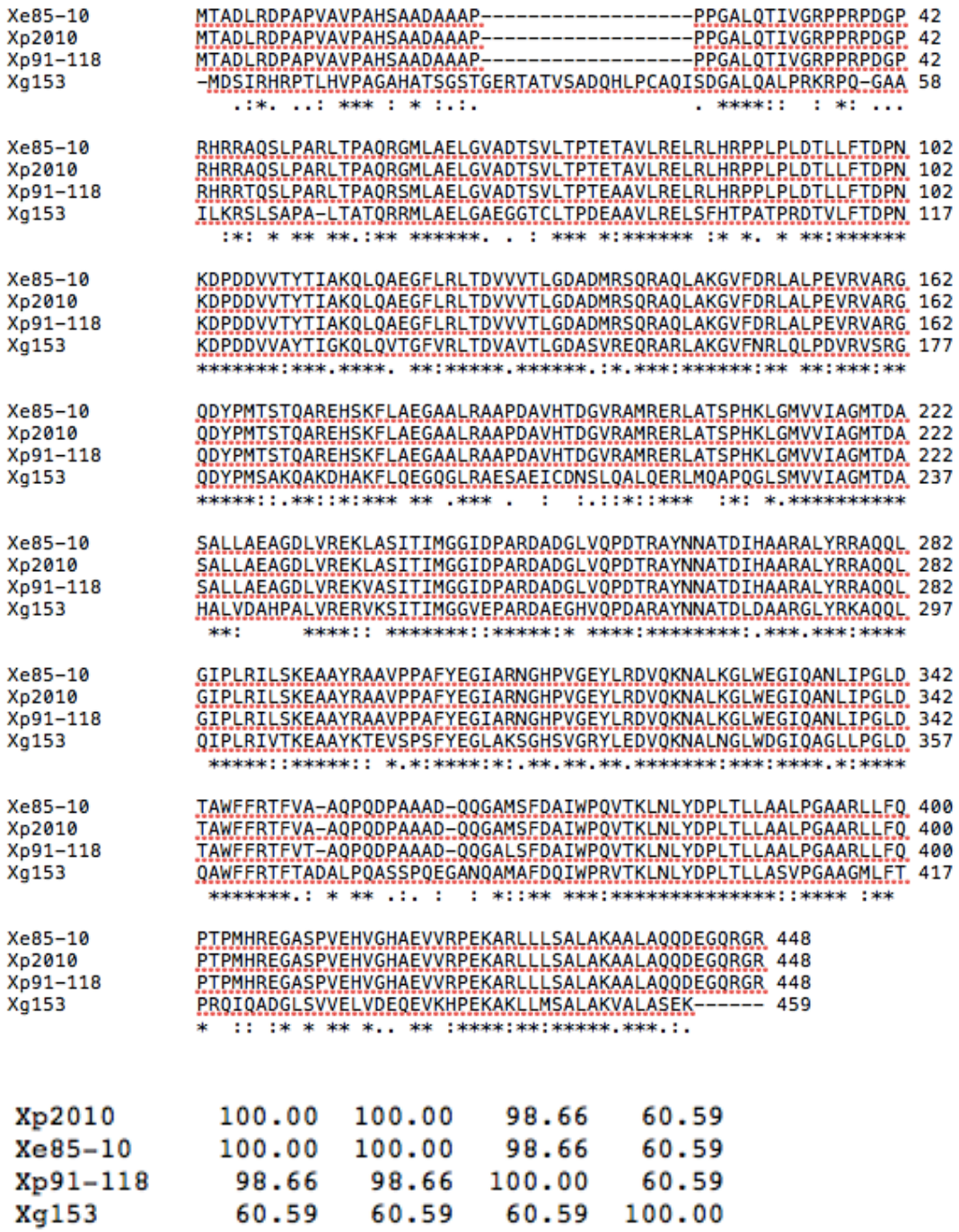


Figure 3-5. Alignment of XopQ alleles from Xe, Xp, and Xg
Clustal Omega alignment of amino acid sequences from XopQ alleles and percent identity matrix.

Discussion

In this section, we presented three examples where the host range of xanthomonad pathogens was expanded by the removal of one or two effectors that tip the balance from avirulence to virulence, summarized in a spectrum of host expansion in Figure 3-7.

In the first example, we took advantage of our whole genome phylogenomics and associated metadata from field strains isolated from diseased pepper and tomato. The majority of Xp strains in our collection, isolated after 1998, have acquired AvrBsT, an avirulence protein responsible for restricting Xp host range on pepper. AvrBsT has been shown to be a virulence factor that suppresses defense responses in tomato (47), possibly conferring a competitive advantage to Xp in tomato fields. However, one field isolate of Xp, Xp2010, was isolated from diseased pepper. Genome sequencing and effector predictions of this strain show that Xp2010 does not possess the AvrBsT. We found that four of the five field strains in our collection (all isolated after 1998) that do not possess AvrBsT are also pathogenic on pepper. We constructed insertion mutants in Xp strains that possess AvrBsT and found that mutation in *avrBsT* results in differences in the *in planta* populations in pepper when compared between Group 1A and Group 2. *avrBsT* mutants in Group 2 experience a full virulence gain on pepper, whereas *avrBsT* mutants in Group 1A acquire only a partial growth benefit, indicating that additional factors restrict the host expansion of Group 1A strains onto pepper. Phenotypic characterization, including pepper pathogenicity tests of *avrBsT* mutants, will need to be conducted on other strains in Groups 1 and 2 to support more definitive conclusions.

In the second example, we tested whether a deletion of *xopQ* in Xe would result in a host gain for *N. benthamiana*. We found that this was indeed the case, mirroring the effect of a hopQ1-1 deletion in Pto DC3000 (48). In the third example, we combined our knowledge from our previous work and hypothesized that a double deletion of XopQ and AvrBsT would expand the host range of Xp for *N. benthamiana*. We found that a double deletion of *xopQ* and *avrBsT* in Xp4B (Xp4BΔXopQΔAvrBsT) results in a *N. benthamiana* host gain. AvrBsT is specific to Xp and has been lost naturally by several strains in our collection, allowing these strains to gain pepper pathogenicity. Whether or not these natural AvrBsT mutants would be as competitive on tomato against Xp with AvrBsT is unknown. Because we have identified AvrBsT as a relatively unstable effector over time in our collection of Xp strains, the R-protein in pepper or *N. benthamiana* that recognizes AvrBsT is not an ideal candidate for genetic resistance to deploy in tomato.

XopQ is the sole avirulence factor restricting Xe pathogenicity on *N. benthamiana*. Because XopQ is a core effector and has been stably maintained since the split of Xe and Xp, it is unlikely that the loss of this effector would occur in nature. Thus, we have identified a source of R-protein in *N. benthamiana* (and *N. tabacum*) that would be a promising candidate as a genetic resistance tool against Xe and Xp in tomato. This "R-XopQ" could be identified using traditional map-based cloning combined with newer methods, such as RenSeq (49). If the "R-XopQ" from *N. benthamiana* is deployed into tomato, it would provide defense against Xe and Xp but not Xg. For example, "R-XopQ" would be useful in protecting tomatoes against bacterial spot in Florida (where the dominant tomato pathogen is Xp), but not in the Midwest (where the dominant tomato pathogen is Xg). Future disease resistance strategies should consider the dominant pathogens in the area

(with their particular effector alleles) and continue to survey and sequence pathogens in the field. This will allow a faster degree of responsiveness to emerging pathogens due to the selection pressures from resistant plants.

Lastly, we identified Xg as a natural pathogen of *N. benthamiana*. Because *N. benthamiana* is developing into a model plant system, the discovery of a natural xanthomonad pathogen will be a useful tool for future studies of plant immunity and plant-pathogen interactions.

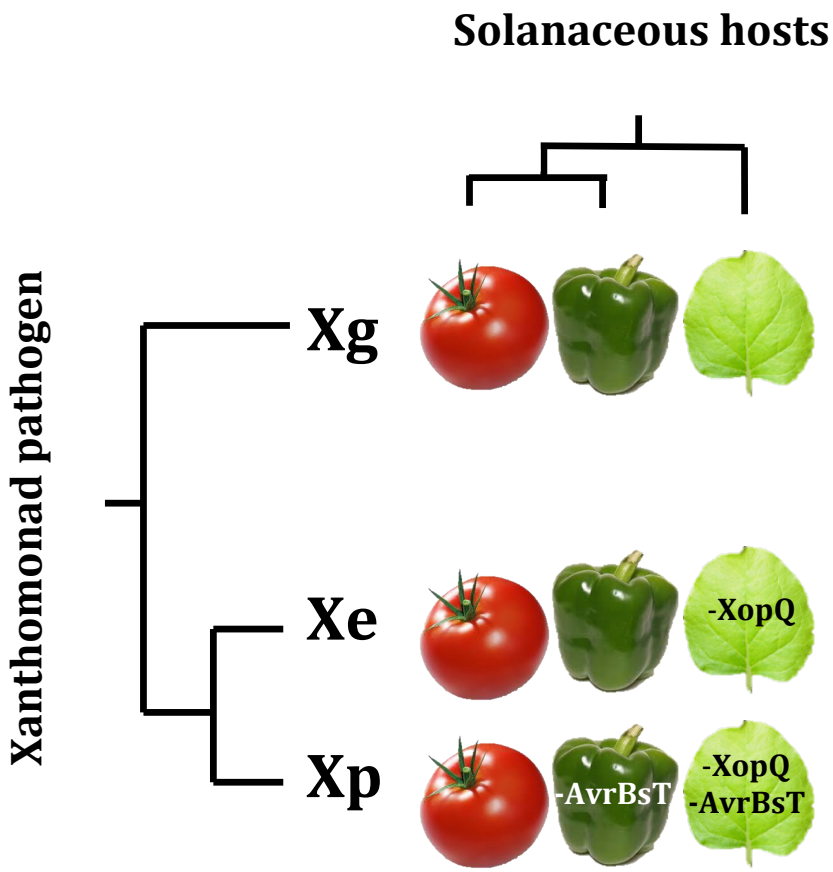


Figure 3-7. Spectrum of solanaceous host of Xe, Xp, and Xg
Xg is a natural pathogen of tomato, pepper, and *N. benthamiana* (plant hosts arranged in order of phylogenetic relatedness). Xe is a natural pathogen of tomato and pepper. A XeΔXopQ mutant causes disease on *N. benthamiana*. Xp is a natural pathogen of tomato. A XpΔAvrBsT mutant causes disease on pepper (although this appears to be limited to strains in Group 2). An Xp double deletion of *xopQ* and *avrBsT* causes disease on *N. benthamiana*. Xg is a natural pathogen of tomato, pepper, and *N. benthamiana*.

4. Characterization of water soaking conferred by AvrHah1, a TAL effector from Xg

Background

Bacterial spot disease caused by *Xanthomonas gardneri* (Xg) has emerged relatively recently as an agronomically important pathogen in the United States. Outbreaks of bacterial spot caused by Xg in Midwestern tomato fields were characterized by severe fruit spotting, which significantly reduces the yield of marketable fruit (13). Although symptoms and lesions are a common effect of plant disease, few studies look into how and why they develop. Severe disease lesions are directly linked to increased yield losses, such as through plant defoliation and sun scorch of fruit. The environment that a plant and pathogen coexist in is a major factor that determines whether or not disease will occur and the degree of symptom severity (5). This relationship is depicted in the classic disease triangle model, where the susceptibility/resistance of the host (plant), the avirulence/virulence of the pathogen, and the complex and dynamic combination of environmental factors all contribute to disease development (50). Because it is the most chaotic and unpredictable edge of the triangle, it has been convenient to ignore the environment by studying molecular plant pathology in a controlled laboratory setting. This reduces variability and increases experimental robustness but does not reflect the conditions a plant will actually face in the field.

The role of water is a particularly important factor in the environment that contributes to the success of plant disease resistance. The most compelling reason to study the role of water in plant-pathogen relationships is that it directly affects the success of the plant's defense response against pathogens (51, 52). Empirical knowledge has maintained that the worst disease outbreaks of bacterial spot occur following or during periods of high humidity and high temperature (100). In the case of foliar bacterial pathogens, plants restrict the flow of water to infected areas in the leaf, thereby restricting the source of nutrients the bacteria depend upon to multiply and spread. This defensive vascular restriction occurs in response to pathogens appropriate without Type III effectors, indicating that it is active at the broad, basal level of plant immunity, PTI. In response to vascular restriction during PTI, several bacterial effector proteins have been shown to prevent the plant's ability to restrict water (53). Water restriction is also important for the success of the ETI response, as relative humidity or high apoplastic water both prevent the formation of HR (52). Furthermore, high temperatures have been shown to block dominant R-gene resistance (54). As the effects of global climate change become more pronounced, the environment must be considered as a major factor in disease resistance strategies.

An opportunity to study the contribution of the environment to symptom development arose with the first characterization of AvrHah1, the single Transcription Activator Like (TAL) effector protein in Xg (see Introduction for additional background on TAL effectors). AvrHah1 was initially identified in a forward screen to find the genetic element that allowed Xg to form water-soaked disease lesions on *bs5* and *bs6* pepper. These pepper varieties carry recessive resistance that is successful against Xe. Xe carrying

cosmid clones from an Xg genomic library were screened on pepper for a gain of water soaking, which led to the identification of AvrHah1 (27). Even though a mutant of Xg without AvrHah1 lost the ability to cause water soaking in pepper, the *in planta* bacterial growth of the mutant did not differ from wild type Xg (27). These results suggest that AvrHah1-induced water-soaked disease lesions are uncoupled from bacterial growth, pointing to a virulence effect operative after apoplastic growth, such as promoting bacterial egression to the leaf surface or transmission to neighboring plants.

We first sought to characterize whether or not AvrHah1 was conserved in the Xg field strains from our xanthomonad collection, specifically those isolated recently in the Midwest. Because the DNA sequences of TAL effectors are highly repetitive, we were not able to assemble AvrHah1 in the genomes of Xg field strains given the short read technology of Illumina sequencing (MiSeq 250 PE) (Table 5). We performed Southern blot analysis on genomic digests of Xg field strains with a probe for AvrHah1. Because AvrHah1 was previously reported to activate Bs3 resistance, we also tested for the activation of a defense response in pepper 30R, a which contains *Bs3*. After confirming that AvrHah1 was a conserved effector (and therefore a potentially good candidate for future plant genetic resistance strategies), we sought to understand how water soaking develops in the tomato and *N. benthamiana*. To do this, we created an XgΔAvrHah1 mutant and characterized the effects of loss of AvrHah1 on water soaking. We observed that AvrHah1 enables the absorption of water into the apoplast of Xg-infected leaves, conferring a dark, water soaked appearance. The AvrHah1-mediated intake of water can be observed in real time under appropriate conditions and can also be measured quantitatively by collection and weighting of apoplastic fluid. Furthermore, bacterial cells can be ferried into the apoplast during water soaking and bacterial surface population is enhanced in Xg wt- compared to XgΔAvrHah1-infected plants. Lesions are the site of cell death and tissue collapse, indicating that water soaking may promote bacterial escape to the cell surface.

Results

Xg field strains contain AvrHah1

Because the repetitive nature of TAL effector genes renders them difficult to assemble from short read DNA sequencing technology (MiSeq 250 PE), Southern blot analysis on genomic DNA was used to detect TAL effector family members (Fig. 4-1A). AvrHah1 is a 3,117bp gene. The majority of the gene (3bp to 2,967bp) is contained within a BamHI fragment. A radiolabeled probe for the first 750bp of AvrHah1 hybridized to a singly band at approximately 3kb, matching the predicted size of the AvrHah1 digested fragment (2,964bp). Although the probe was amplified from an AvrHah1 template, the high sequence homology of TAL effectors—even a relatively divergent TAL effector like AvrHah1—could hybridize to other TAL effectors. Cloning and sequencing of the TAL effector from five randomly selected Xg field strains revealed that the TAL effector was indeed AvrHah1. Xg field strans for Ohio and Michigan strains are labeled Xg152-183. Xg137 (Xg 1782) and Xg140 (04T5) serve as negative and positive controls, respectively, for AvrHah1 (27).

Bs3 is an executor (E) type of resistance gene that possesses a promoter EBE "trap" and is activated only in response to corresponding TAL effectors (55). Because Bs3 resistance, present in pepper cv. 30R, is activated by both AvrBs3 (56) and AvrHah1 (27), Xg field strains were infiltrated into pepper cv. 30R to confirm TAL effector activity (Fig. 4-1B). All Xg field strains (Xg152-Xg183) activated Bs3 resistance, but not strains Xg137 and Xg138, which do not possess AvrHah1 and do not active Bs3 resistance. As expected, all Xg strains activate a HR on pepper 20R, which possesses the Bs2 resistance gene (57) that activates an HR in response to the cognate effector AvrBs2.

We found that AvrHah1 was present and active in all Midwestern Xg field strains from our collection. Because we found that AvrHah1 is a conserved effector in Xg, genome-editing strategies against AvrHah1 target EBEs could be used to increased disease tolerance in plants. Stacking resistance genes like Bs2 and Bs3 with mutations in AvrHah1 target EBEs could potentially increase both durability and resistance.

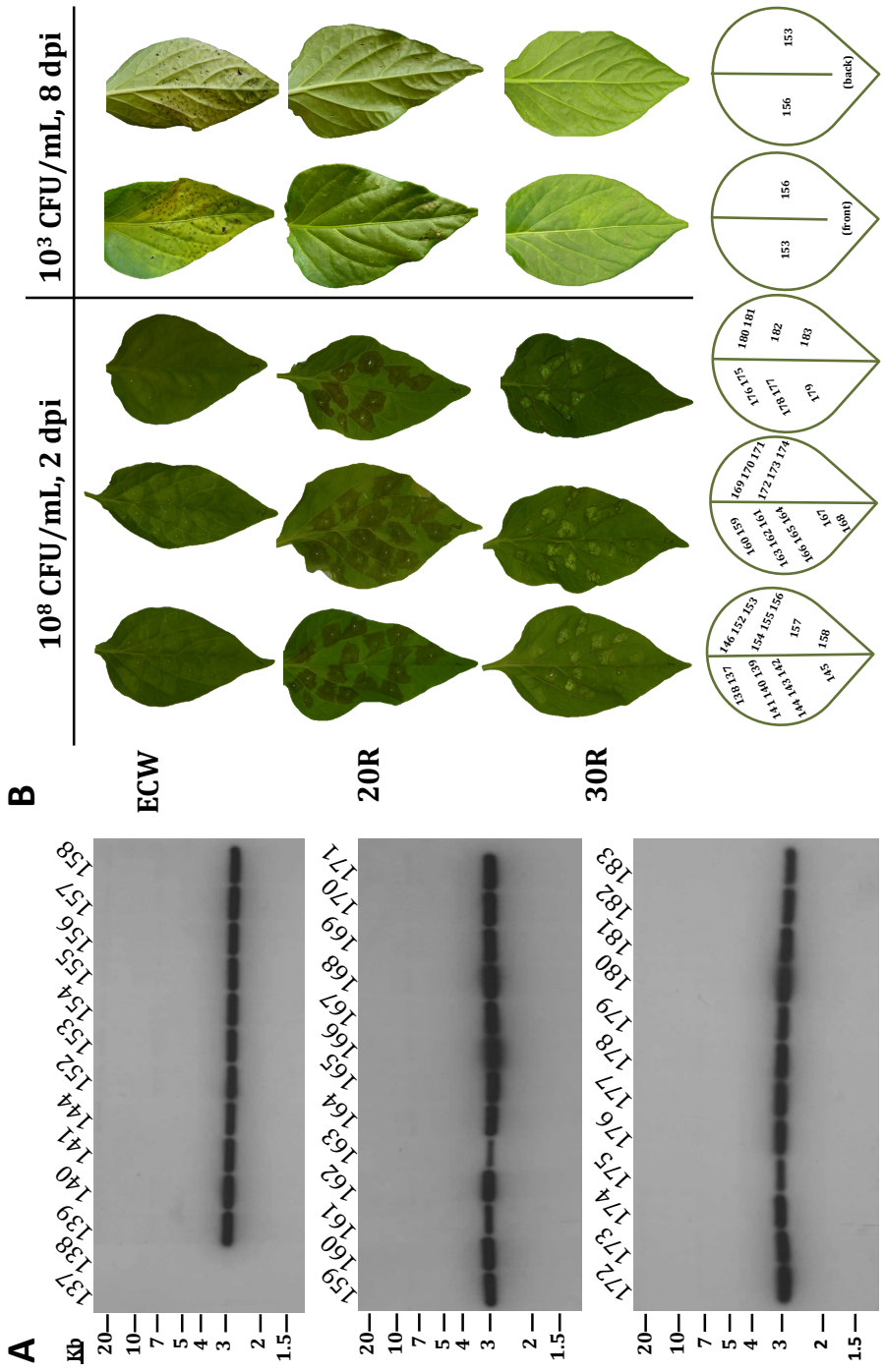


Figure 4-1. All Xg field strains contain AvrHah1, a 3kb TAL effector that activates Bs3 resistance

(A) Southern blot analysis of Xg genomic DNA using a probe for the first 705bp of AvrHah1. Field strains from the Midwest are numbered 152-183. The size of the predicted BamHI-digested AvrHah1 fragment is 2,964bp. Negative controls and positive controls for AvrHah1 in Xg are strain 1782 and 04T5, respectively. (B) All Xg field strains activate Bs2 (20R) and Bs3 (30R) resistance and are pathogenic on pepper ECW

Construction of XgΔAvrHah1

Because AvrHah1 could not be assembled from our high throughout sequencing data, it was not easy to determine upstream and downstream sequences to create a whole gene knockout using a double homologous recombination strategy. Primers with XbaI sites were designed to PCR amplify the sequences flanking the central DBD for ligation (5', 1-805-XbaI bp and 3', XbaI-2280-3117bp) in order to create an in-frame deletion of the ~1.5kb of the central DNA Binding Domain (Fig. 4-3), or AvrHah1ΔDBD. This mutant is referred to as XgΔAvrHah1, as it is a functionless version of the effector. This sequence was cloned into the suicide vector pLVC18 *sacBR* Tet^R (58). After conjugation of this vector into Xg153 and selection of tetracycline resistance and sucrose sensitivity, exconjugants (with integrated pLCV18 construct) were cycled on rif to allow for a second recombination event that replaces the wild type AvrHah1 with the mutant AvrHah1ΔDBD. True knockouts were markerless (Tet sensitive). Whole gene knockouts of *hrcV* and *avrBs2* and a double mutant in *avrBs2* and *avrHah1* were also similarly constructed in Xg153 to create an isogenic line of mutants. HrcV is a major structural component of the Type III secretion system and required for secretion of effectors into plant cells (59). AvrBs2, as previously mentioned, is a core effector present in most xanthomonads that is required for full pathogen virulence (101).

To confirm knockout mutations, 5 micrograms of BamHI digested genomic DNA from Xg153 wt and XgΔhrcV, XgΔAvrBs2, XgΔAvrHah1, and XgΔAvrBs2/ΔAvrHah1 mutants was analyzed via Southern blot (Fig. 4-3). DIG-labeled probes against the entire gene length of *hrcV* and *avrBs2* were used and hybridized to bands at 7kb and 10kb, respectively, in the presence of the gene but not in the corresponding mutants (Fig. 4-3A, B).

The size of BamHI-digested AvrHah1 is 2,964bp and the size of the BamHI-digested AvrHah1ΔDBD deletion is 1,491bp. A probe for the central DBD (768bp-2355bp) hybridizes to a 3kb band only in the presence of full length AvrHah1 (Fig. 4-3C). A probe for the N-terminal fragment of AvrHah1 (1bp-805bp) hybridizes to a 3kb band when full length AvrHah1 is present and to a roughly 1.5kb band in the case of AvrHah1ΔDBD (Fig. 4-3D). Additionally, XgΔAvrHah displays a 1oss of HR on Bs3 pepper (Fig. 4-4), indicating that the DBD deletion is a true non-functional mutation in AvrHah1. A GFP tag was translationally fused upstream of wt AvrHah1 and AvrHah1ΔDBD and cloned into a binary expression vector for Agrobacterium delivery. Western blot analysis indicates that the expressed AvrHah1ΔDBD is indeed an in-frame deletion protein at the expected size (Fig. 4-5).

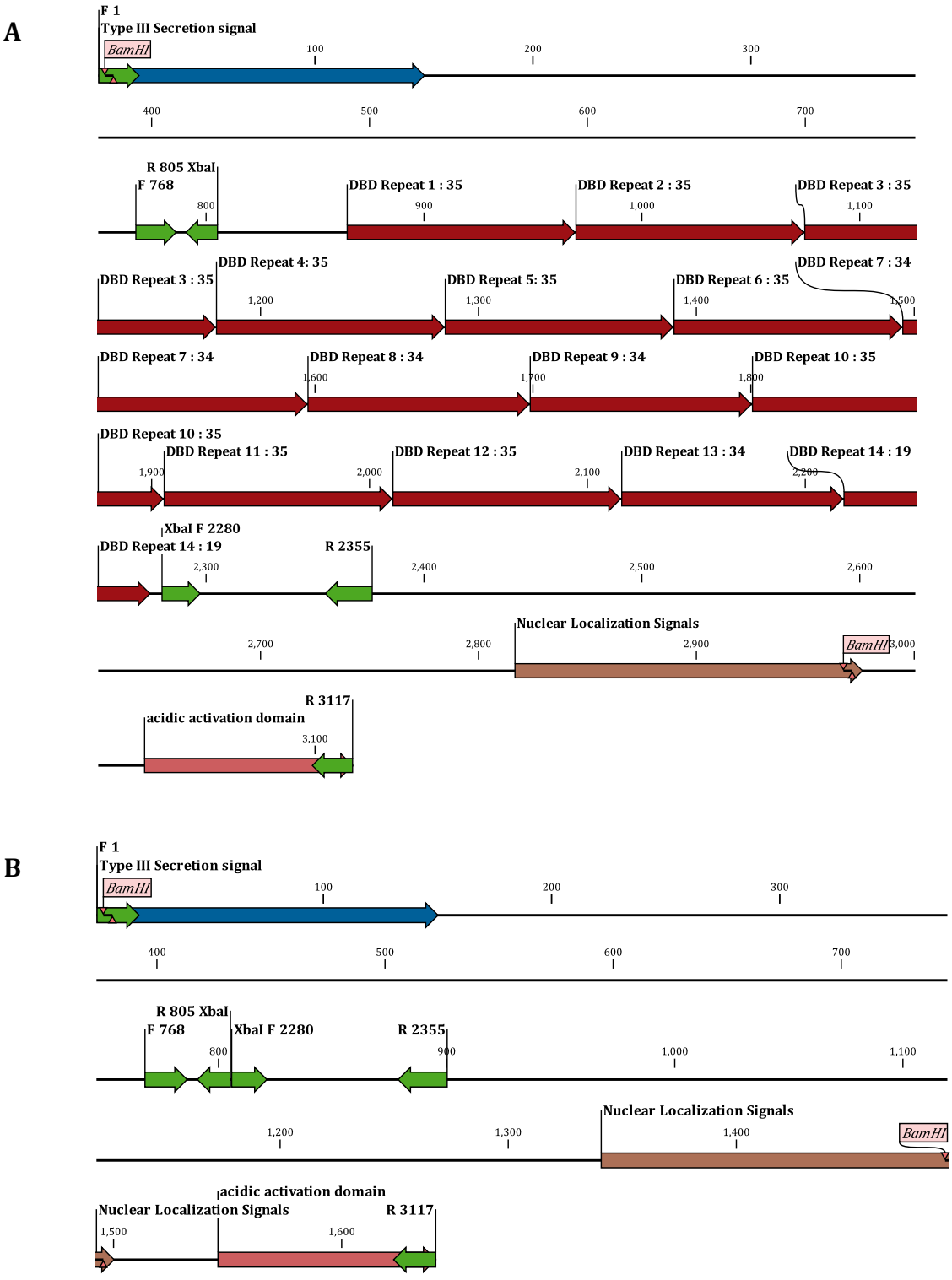


Figure 4-2. Gene map of AvrHah1 and AvrHah1ΔDBD
(A) Full length gene map of AvrHah1 showing relevant primer binding sites (green), including XbaI sites for cloning around the central DBD (red arrows, each repeat). BamHI sites used for Southern blot analysis are also shown. (B) The in-frame deletion of the DBD.

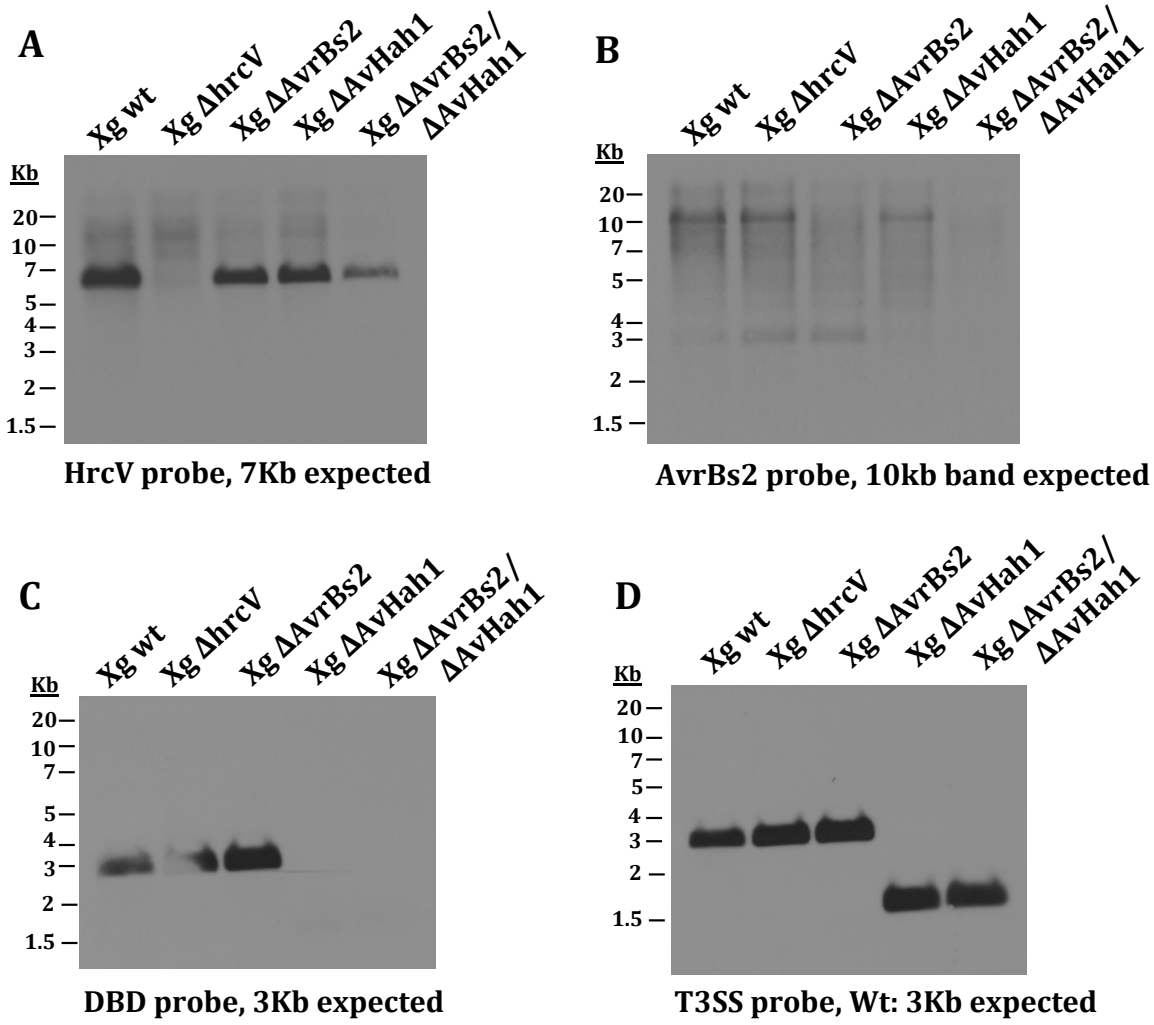


Figure 4-3. Confirmation of mutants in Xg153
Southern blot analysis of Xg deletion mutants. The same membrane was stripped and re-probed for (A) HrcV, (B) AvrBs2, (C) the DBD of AvrHah1 (768bp-2355bp), and (D) the first 705bp of AvrHah1.

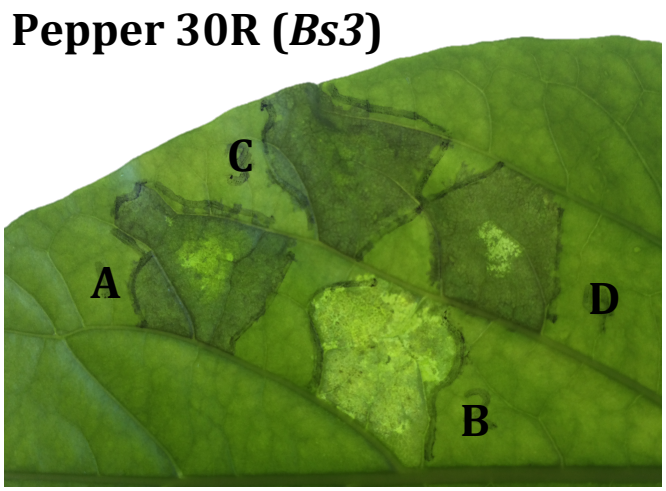


Figure 4-4. AvrHah1ΔDBD is non-functional
Pepper 30R was infiltrated with A.) Xg wt, B.) XgΔAvrHah1, C.) XgΔAvrHah1 + AvrHah1, D.) XgΔAvrHah1 + AvrBs3. (48hpi, OD_{600} = 0.1). A dark cell death (HR) in A, C, and D is from *Bs3* activation.

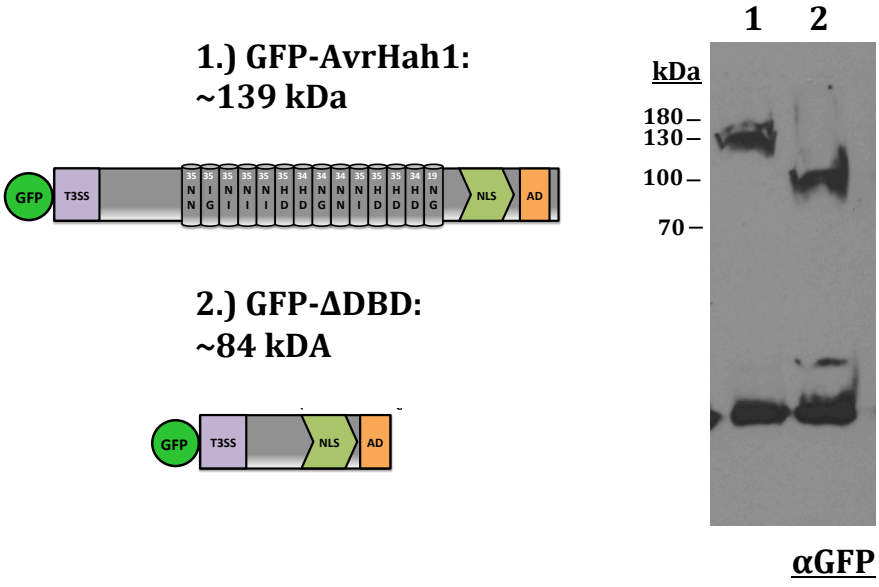


Figure 4-5. AvrHah1ΔDBD produces an in-frame protein
GFP was cloned N-terminal to (1) AvrHah1 and (2) AvrHah1ΔDBD. *Agrobacterium* transient expression in *Nicotiana benthamiana* and western blotting for GFP show their expression and relative protein sizes (OD_{600} = 0.3, 24 hpi).

Characterization of AvrHah1-dependent water soaking

We observed a dark, apoplastic water soaking when Xg, but not XgΔAvrHah1, was infiltrated into tomato, pepper, and *Nicotiana benthamiana*, the known hosts of Xg (60) (Fig. 4-6.). Because we were interested in the recent Xg bacterial spot outbreaks in the Midwest, we investigated the role of AvrHah1 in tomato water soaking. Similar to previous results in pepper, we found that Xe85-10 carrying AvrHah1 is able to induce water soaking in tomato (Fig. 4-7). Xe alone induces dry, flecked lesions on tomato. Xe carrying AvrHah1ΔDBD induces lesions similar to Xe alone. Xe carrying AvrBs3 does not induce disease lesions, likely due to recognition by the tomato Bs4 R protein. The differential recognition of AvrHah1 and AvrBs3 will be discussed Section 6.

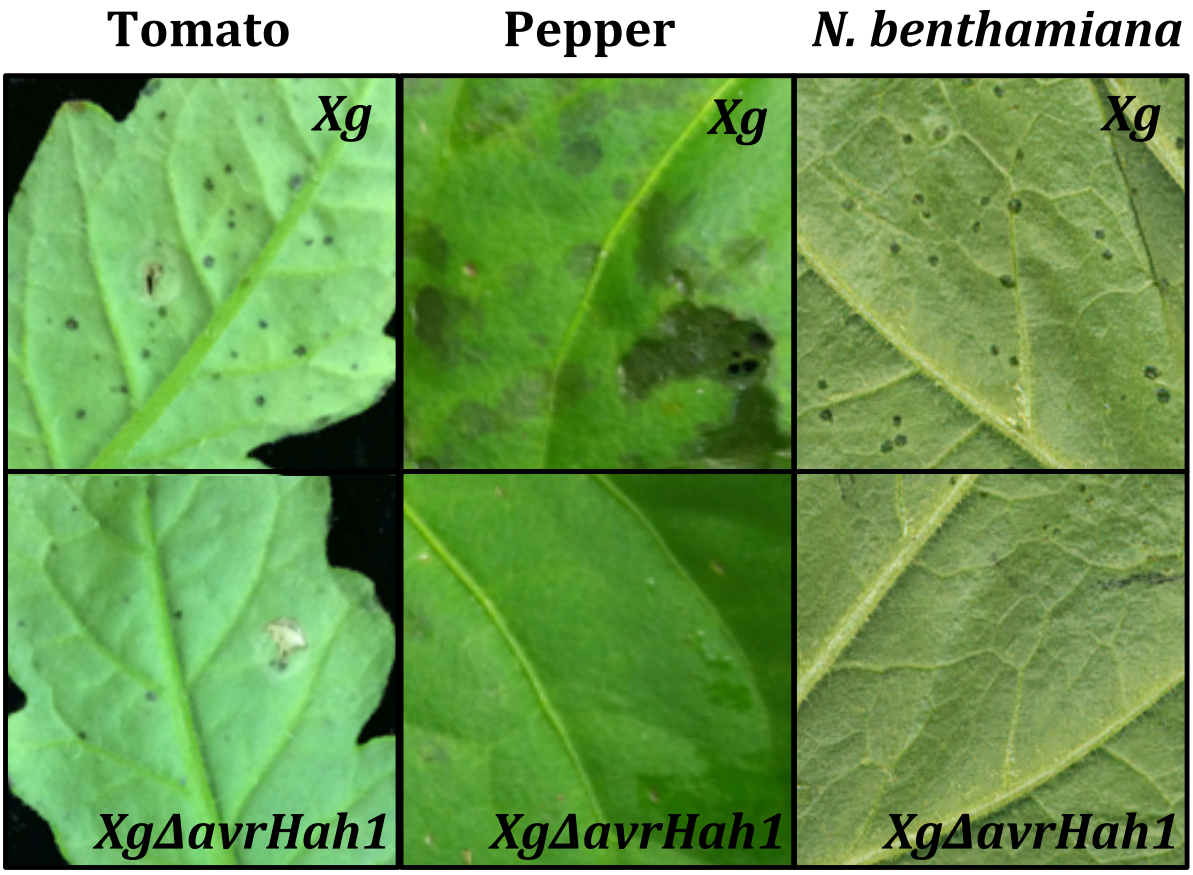


Figure 4-6. XgΔAvrHah1 has reduced water soaking on known plant hosts
Leaves were syringe infiltrated at 10^4 CFU/mL and photographed 6-8 dpi.

Although water-soaked symptoms develop at ambient conditions (such as in Fig. 4-7), we observed enhanced water-soaked symptoms when infected plants were placed in a mist chamber. We could also obtain enhanced water soaking by growing infected plants at ambient conditions and then submerging leaves in water. Remarkably, we observed the uptake of water into the leaf apoplast in real time when Xg carrying AvrHah1 was infiltrated into *N. benthamiana* (Fig. 4-8). At 48hpi, a 30 ul drop of water was pipetted on top of a slight epidermal wound in the circled infiltrated areas. The drop of water on the leaf surface shrunk as it was pulled into the leaf, darkening the apoplast. The darkened, advancing front proceeded away from the wound and stopped at the edge of the infiltrated area. By two minutes, the zones infiltrated with Xg or XgΔAvrHah1 + AvrHah1 was completely darkened by water soaking, while the water drop remained on top of the leaf where XgΔhrcV or XgΔAvrHah1 are infiltrated.

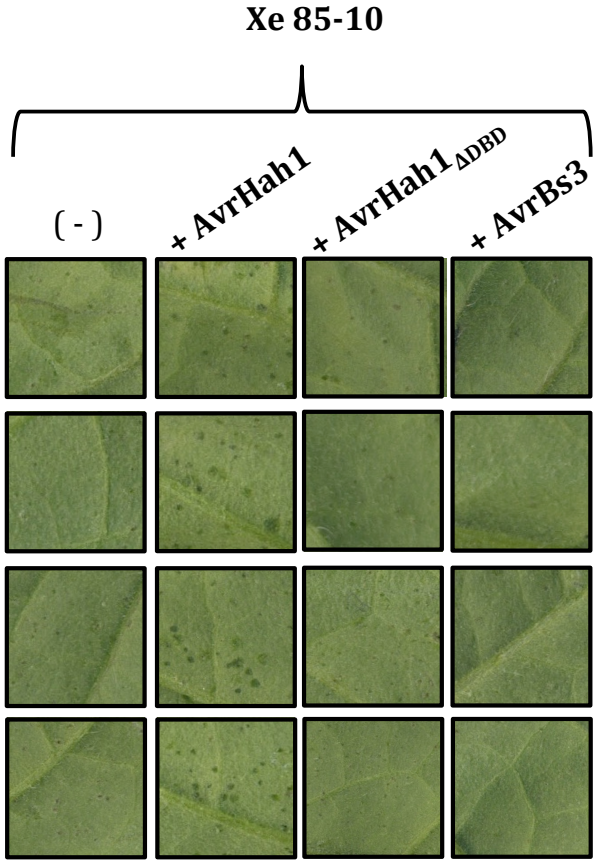


Figure 4-7. AvrHah1 confers water soaking to Xe in tomato
Tomato Heinz 1706 was infiltrated at a low inoculum of 10^4 CFU/mL (for development of single lesions) with Xe85-10 alone (-) and Xe carrying AvrHah1, Avrhah1$_{\Delta DBD}$, and AvrBs3. At 6dpi, Xe + AvrHah1 develops water soaked disease lesions, while Xe alone and Xe + AvrHah1$_{\Delta DBD}$ develops flecked lesions. Xe + AvrBs3 does not develop any lesions, due to the recognition of AvrBs3 by the tomato R protein Bs4.

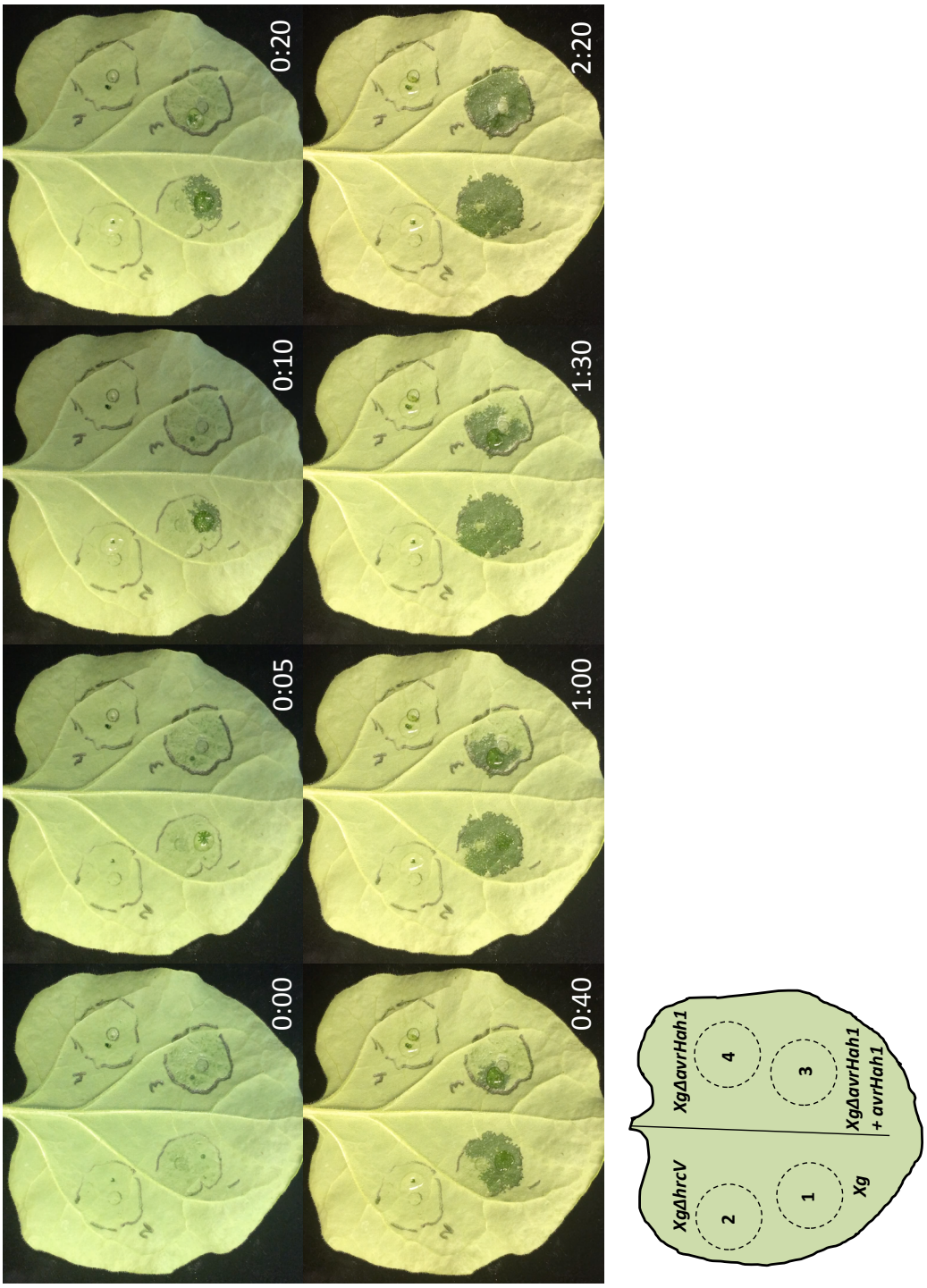


Figure 4-8. AvrHah1 promotes the intake of water into the apoplast of Xg-infected plants in *N. benthamiana*

Nicotiana benthamiana was syringe infiltrated (circled areas) with 1.) Xg wt, 2.) XgΔhrcV, 3.) XgΔAvrhah1 + AvrHah1, 4.) XgΔAvrhah1 (OD_{600} = 0.1) ([1]bottom left, [2]top left, [3]bottom right, [4]top right, respectively, as indicated by diagram). At 48hpi, a 30 ul drop of water was pipetted on top of a slight epidermal wound in the circled infiltrated area.

In tomato, the water soaking effect of Xg was similarly dramatic and enhanced by external water, yet progressed more slowly (20 minutes compared to 2 minutes in *N. benthamiana*) (Fig. 4-9). Xg shows a typical water soaking response, which is absent in XgΔhrcV and XgΔAvrHah1. Complementation of XgΔAvrHah1 + AvrHah1, but not AvrHah1ΔDBD, restores water soaking. AvrBs3 induces a cell death, or Hypersensitive Response (HR), due to the recognition of AvrBs3 by the Bs4 R protein in tomato.

To quantitate the water soaking effect, we collected (by centrifugation) and weighted the apoplastic fluid from Xg-infiltrated tomato leaves that were submerged in water for 20 minutes (Fig. 4-10). For comparison, we also measured the apoplastic fluid from leaves without any Xg ('untreated') and leaves that were infiltrated with water immediately prior to tissue collection ('infiltrated'). There is approximately five times more apoplastic water present in leaves fully infiltrated than leaves with no water soaking. In the presence of wt Xg there is about a six fold increase in apoplastic water compared to untreated leaves and leaves infiltrated with XgΔAvrHah1 or XgΔAvrHah1 + AvrBs3. Complementation of AvrHah1 into XgΔAvrHah1 restored water soaking to wt levels.

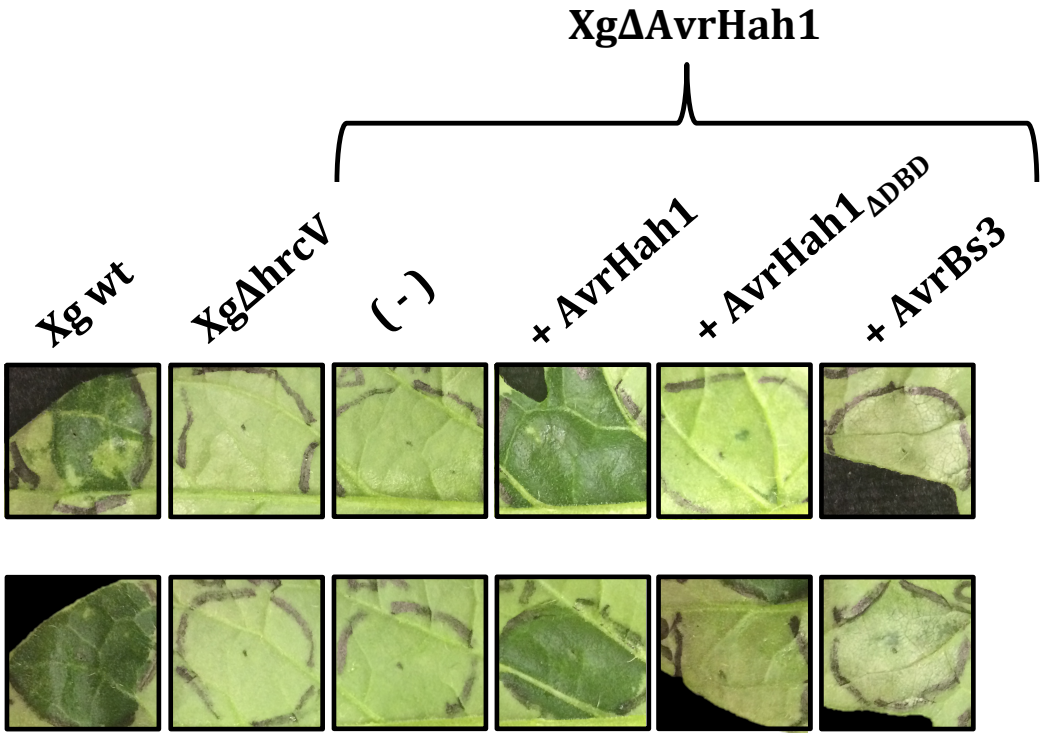


Figure 4-9. AvrHah1 promotes the intake of water into the apoplast of Xg-infected plants in tomato
Heinz 1706 was infiltrated at a high inoculum OD_{600} = 0.1 (~10^8 CFU/mL). At 48 hpi, leaves were submerged in water for 20 min and photographed (2 replicates are shown).

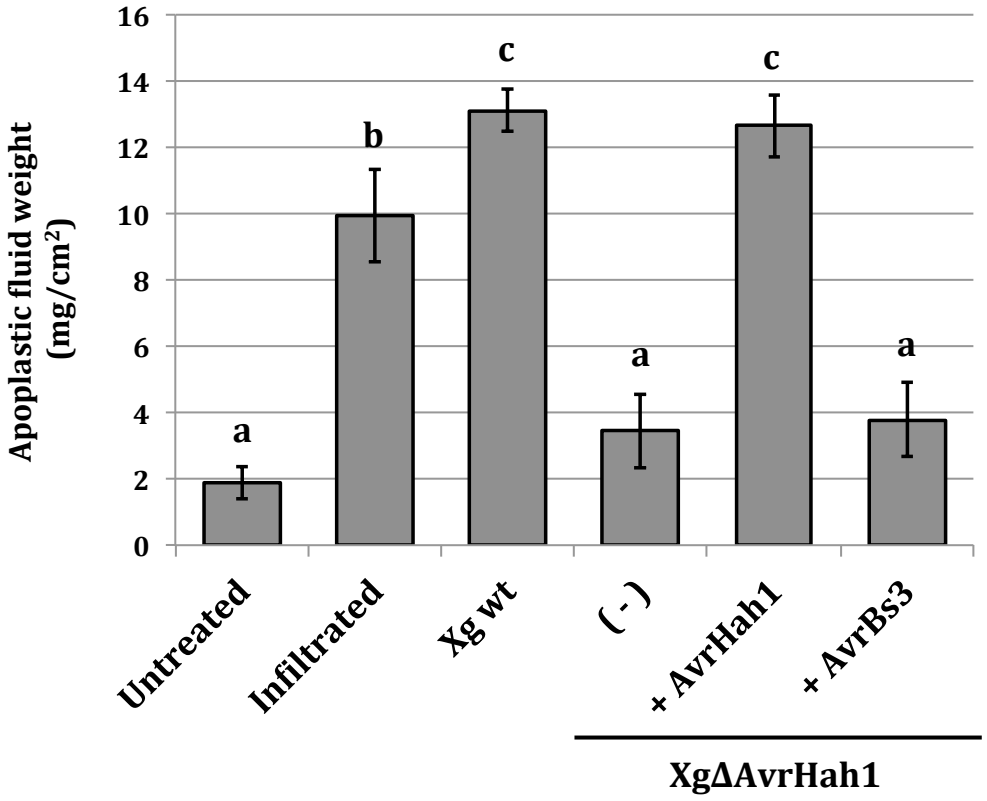


Figure 4-10. Quantitative assay to measure water soaking from AvrHah1
Heinz 1706 was infiltrated at a high inoculum OD_{600} = 0.1 (~10^8 CFU/mL). At 48 hpi (at which point leaves did not show any obvious water soaking), leaves were submerged in water for 20 min and the apoplastic fluid was collected by centrifugation. An unpaired t-test compared the mean weight from each condition, different letters above the bars denote statistical difference ($p < 0.5$, n = 8 samples), and standard error bars are shown.

Previous work has shown that AvrHah1 does not contribute to apoplastic growth in pepper. We found similar results in tomato (Fig. 4-11). We did not detect a growth defect for AvrHah1 in the background of a XgΔAvrBs2 mutant. AvrBs2 is an effector which has a significant growth defect in many xanthomonads including Xe (101) and Xg. Although no differences in apoplastic growth were observed, we detected an increase in bacterial population on the leaf surface of tomatoes infected with Xg carrying wt AvrHah1 compared to AvrHah1ΔDBD (Fig. 4-12).

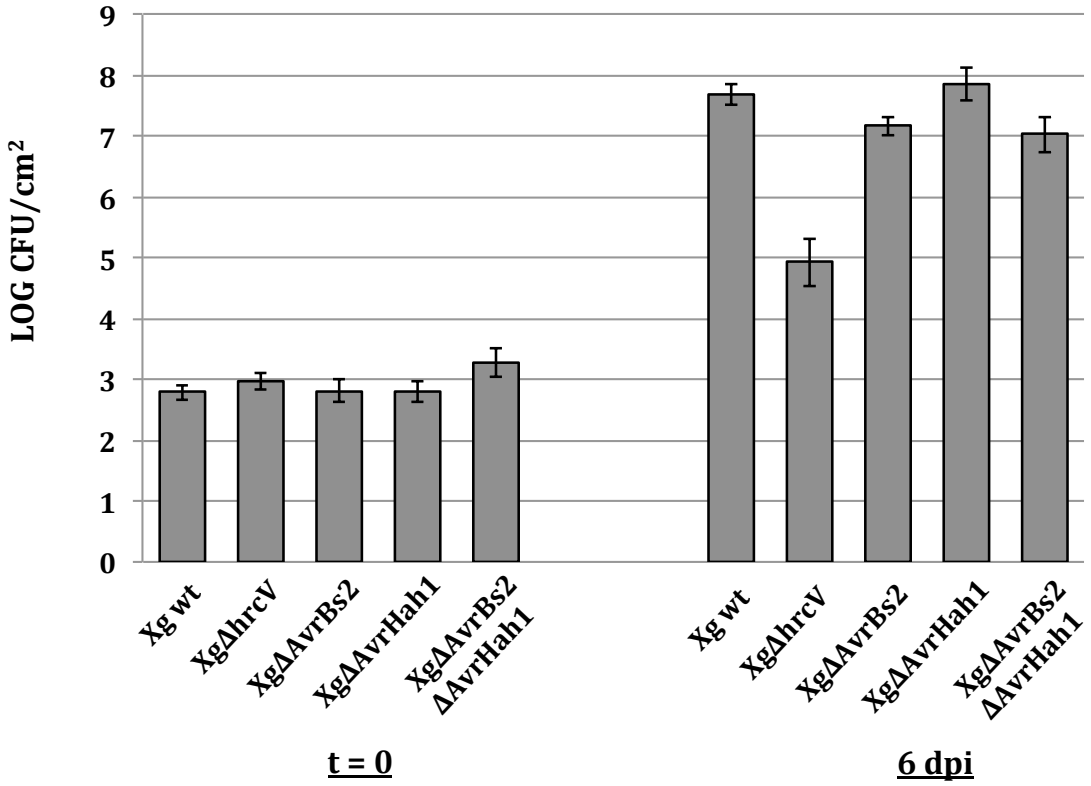


Figure 4-11. AvrHah1 does not promote *in planta* growth in tomato
Tomato Heinz 1706 was infiltrated with Xg strains at 10^4 CFU/mL. Six days post infiltration, the *in planta* growth of bacteria was measured (n = 6 leaf discs). Xg with a double deletion of AvrBs2 and AvrHah1 does not display an additional growth defect compared to a single mutant of AvrBs2.

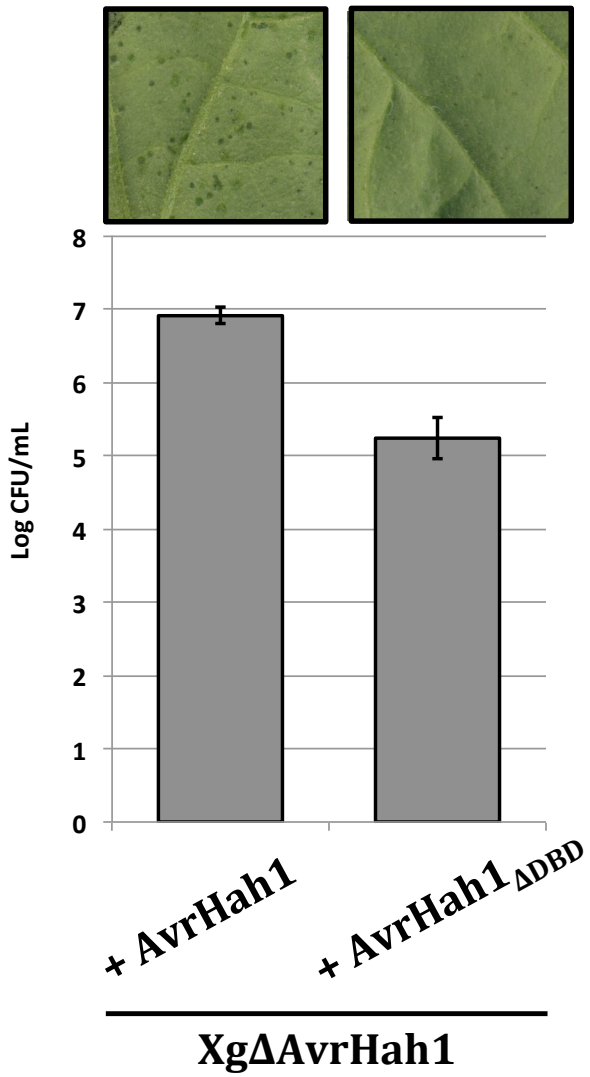


Figure 4-12. AvrHah1 increases bacterial surface population in tomato
Xg strains were infected into tomato at 10^4 CFU/mL, at 8dpi single lesions were apparent. Three 15ul drops of water were allowed to rest on the leaf surface for 20 min and were pooled into one sample prior to dilution and plating. (n = 6 samples from three leaves).

We wondered if water soaking could be a mechanism that introduces surface bacteria into the apoplast. We set up a water soaking assay in *N. benthamiana* with Xg wt and XgΔhrcV as previously described, except instead of a drop of water on the wound site we used a 30 µl drop of dilute Xg with tetracycline resistance (Xg TetR). We plated a dilution series from macerated leaf discs on rifampicin (Rif) to measure all Xg, or rifampicin and tetracycline (Rif +Tet) to select for newly introduced Xg TetR from the water soaking drop (Fig. 4-13). We sampled leaf discs away from the wound site to avoid any Xg TetR cells left on the leaf surface. We found that water soaking could enable the newly introduced tetracycline resistant Xg to enter the apoplast away from the initial wound site. No tetracycline resistant Xg were detected in the apoplasts of leaves infected with XgΔhrcV, which were not water soaked.

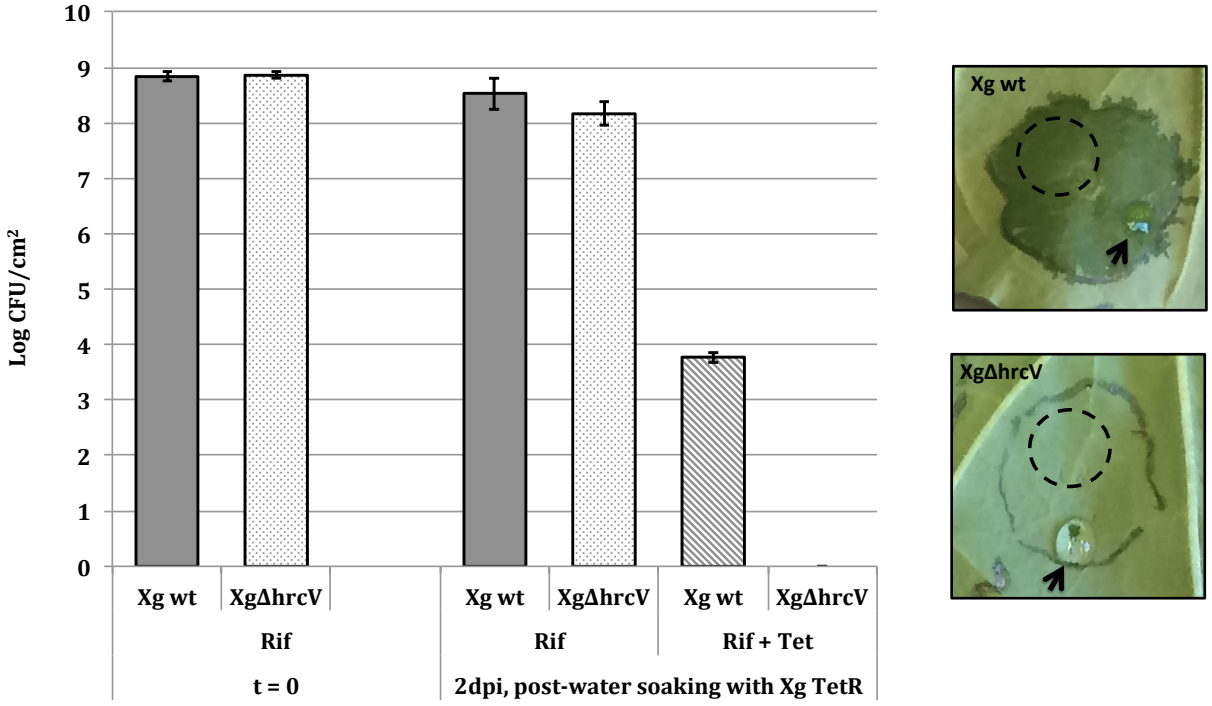


Figure 4-13. Water soaking in *N. benthamiana* encourages intake of external bacteria
N. benthamiana was syringe infiltrated with either Xg wt or XgΔhrcV at OD_{600} = 0.1. Water soaking was induced at 48 hpi with a 30 µl drop of tetracycline resistant Xg (Xg TetR, 10^5 CFU/mL in 10mM $MgCl_2$) on a wound (arrow). Leaf discs were collected away from the wound site after 5 minutes (dashed circle), ground in 10mM $MgCl_2$, and dilutions were plated on either Rif or Rif + Tetracycline to select for the growth of all Xg or the Xg TetR from the water soaking inoculum, respectively (n = 6).

The tissue damage that occurs during infection of Xg is apparent when observing disease lesions on GFP-*N. benthamiana*. This line of *N. benthamiana* stably expresses GFP, which can be visualized under the microscope when excited with blue light (RL). The GFP appears in the cytoplasm of the puzzle-piece like pavement cells of the epidermis, which runs along to the edges of the cell wall due to the large central vacuole (Fig. 4-14). Xg wt and XgΔhrcV were infiltrated into GFP-*N. benthamiana* at a low inoculum (10^4 CFU/mL) to allow for development of discrete lesions. We observed a "dead zone" without GFP signal in the middle of Xg lesions (Fig. 4-14A). In contrast, XgΔhrcV does not form lesions (due to activation of PTI) and the entire surface of the leaf disc expresses GFP (Fig. 4-14B). When observed at greater magnification under normal light (TL), some Xg wt lesions appear to have a physical hole running through its center, as shown by the presence of saturated white light permeating through the leaf (Fig. 4-14C). Under RL to show GFP fluorescence, the central dark zone indicates that cells in the lesion are dead and not expressing GFP (Fig. 4-14E). The increased brightness around the perimeter of the lesion could be because the lesion is itself raised closer to the light source (and appears brighter). When XgΔhrcV is inoculated into GFP-*N. benthamiana*, the tissue is indistinguishable from uninoculated plants. All of the epidermal pavement cells in the XgΔhrcV-infected plants continue to show fluorescence under RL, indicating that they are alive and expressing GFP (Fig. 4-14D, F).

Scanning electron micrographs of Xg wt-infected leaf surfaces show clear disruptions in the epidermal layer. Microlesions are apparent in diseased tissue as single cells or as large areas of disease (Fig. 4-15A, B). High magnification images of single cells show that single epidermal cells appear to have completely collapsed or are in the process of collapse (Fig. 4-15C, D).

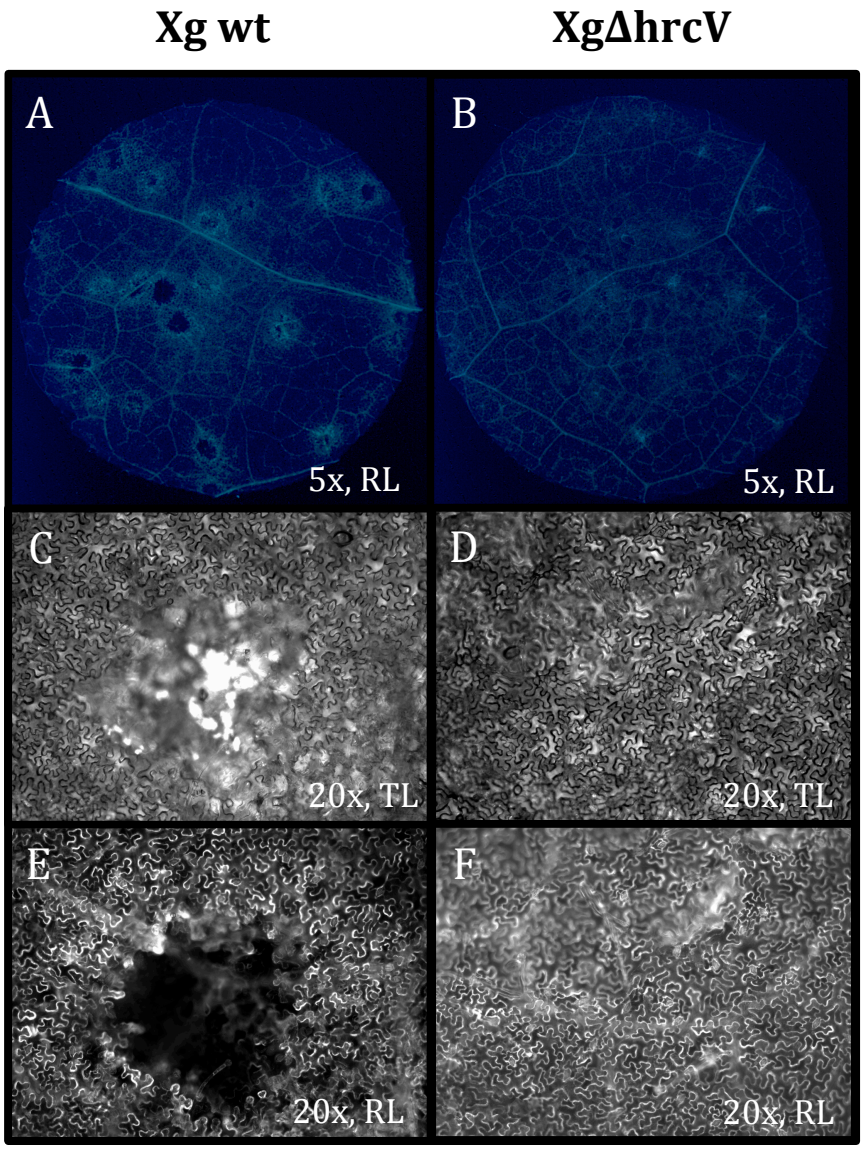


Figure 4-14. Xg lesions are sites of cell death visualized in GFP-*N. benthamiana*
Observations of Xg wt (left column) and XgΔhrcV (right column) on GFP-*Nicotiana benthamiana* at infiltrated at low concentration (10^4 CFU/mL, 8 dpi). Wt Xg forms single disase lesions, while XgΔhrcV appears uninfected due to the activation of PTI. RL = Reflected Light (blue light at 535nm to observe GFP), TL = Transmitted light. (A) A 0.5cm^2 punch of Xg wt-inoculated GFP-*N. benthamiana* was observed under fluorescence at 5X magnification. Single disease lesions are clearly visible as bright rings surrounding a "dead zone" with no fluorescence. (B) A 0.5cm^2 punch of XgΔhrcV-infected GFP-*N. benthamiana*. No lesions are visible. (C) A single lesion magnified at 20X from Xt Xg in normal Transmitted Light, shows a disruption at the cell surface and a shot-hole effect transversing the leaf. (D) The surface of XgΔhrcV-infected area does not show obvious differences from the surface of an uninoculated leaf. (E) The single lesion from (C) visualized for GFP shows the "dead zone" of the lesion. (F) The epidermal cells from (D) are all alive and expressing GFP.

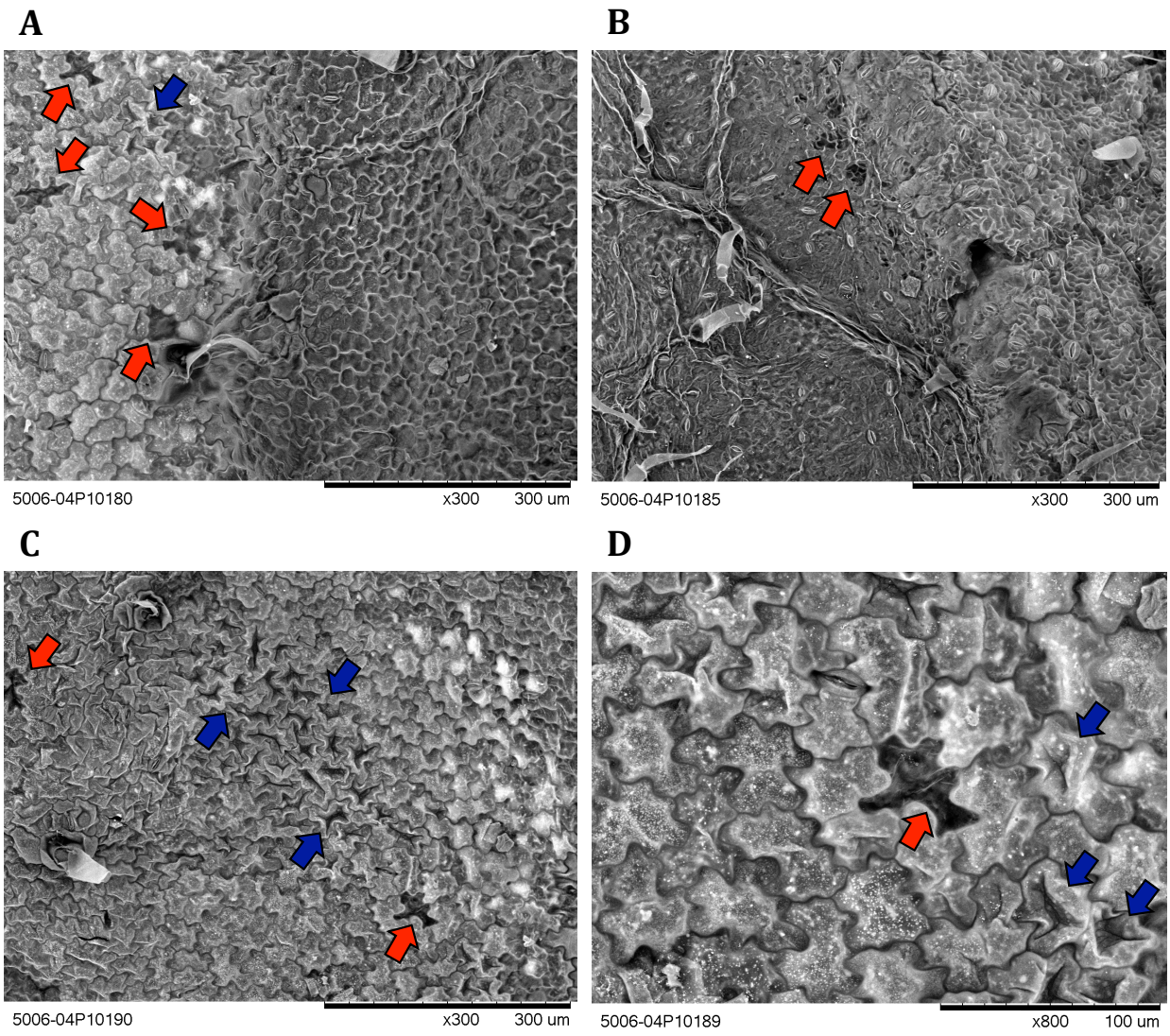


Figure 4-15. Scanning electron micrographs of Xg-infected tomato leaves
Samples were processed by the UC Berkeley Electron Microscope Lab. Images were taken using a Hitachi TM-100 SEM. Red arrows indicate completely collapsed epidermal cells, blue arrows indicate partially collapsed cells. (A) Healthy and diseased tissue are on the left and right of the image, respectively. The border contains cells in partial or complete collapse. (B) Visible holes through the collapsed epidermis are present in the diseased side (left). (C) Large area of partially collapsed cells (depicted by three blue arrows). (D) High magnification bottom right corner from image in (C) depicting completely and partially collapsed cells.

Discussion

Disease outbreaks of bacterial spot are favored in periods of high humidity (100). Given our initial observation that the water soaking caused by Xg was enhanced when the infected plants were placed in a mist chamber, we wondered how water external to the leaf was exacerbating Xg/AvrHah1-induced water soaking. In an agricultural setting, the external water could be from high humidity, rain, or sprinkler (overhead) watering systems. In the laboratory, we chose to submerge leaves in water prior to observation. We could quantitatively measure the intake of water by collecting and weighing the apoplastic fluid after submerging infected leaves in a water bath. We observed that Xg wt, but not XgΔAvrHah1, induced water intake at approximately six times the level measured from untreated leaves (also submerged in a water bath). AvrHah1 was complemented XgΔAvrHah1 back to Xg wt levels. Importantly, the non-functional AvrHah1ΔDBD did not restore water soaking, indicating that the DNA binding function (and gene activation) function of AvrHah1 is responsible for water soaking. AvrBs3 induced a resistance response, due to the activity of the tomato Bs4 R protein, indicating that the water soaking effect is not a result of a resistance response or cell death from HR (the differential recognition between AvrHah1 and AvrBs3 by Bs4 will be further described in Section 6). These results suggest that AvrHah1 may be priming the *in planta* environment such that Xg can take advantage of sudden appearances of external water.

Because there is no difference in apoplastic growth between Xg and XgΔAvrHah1, even in the background of a less virulent mutant (XgΔAvrBs2), the role of AvrHah1 is likely to benefit the bacteria after the apoplastic growth phase of the pathogen's life cycle, such as during bacterial egression to the leaf surface or transmission to neighboring plants. Further tests will need to be performed to determine the roll of AvrHah1-induced water soaking on pathogen fitness and the epiphytic growth phase of the pathogen.

Tissue necrosis (cell death) was observed in the center of lesions from wt Xg on GFP-*N. benthamiana*. It may seem counterintuitive that Xg would be strategically inducing cell death, as this is a strategy also employed by the plant during ETI. However, the typical cell death related to ETI is fast and results in bacterial cell death. The necrosis induced by Xg is likely at such a late stage in infection that bacterial growth has successfully occurred and the physical breaking of the epidermis allows bacteria to escape from the apoplast to the leaf surface. Because Xg is able to grow to high levels in *N. benthamiana* (Fig. 3-4), it is unlikely that the necrosis induced during disease has a toll on overall bacterial fitness. Once on the leaf surface, bacteria would be available to spread to neighboring plants, thus increasing the probability of transmission and re-infection. We observed an increase in bacterial populations on the leaf surface when XgΔAvrHah1 delivered AvrHah1 compared to AvrHah1$_{\Delta DBD}$. We observed an increased degree of leaf tissue damage after water soaking occurred, consistent with the hypothesis that AvrHah1 and water soaking, or the resulting leaf damage from water soaking, aid in the egress of bacteria from the apoplast to the leaf surface. Higher bacterial numbers on the leaf surface, especially during a period of rainfall, would increase the probability of inoculum transmission to neighboring hosts.

The absorption of water into the leaf apoplast infected with Xg is a striking phenotype. We designed an experiment to show that bacterial cells can "hitch a ride" into the apoplast through the mechanism of water soaking. Whether or not bacteria such as

human pathogens can survive in plant apoplast after entry into the leaf from water soaking will need to be determined. This is particularly important in the light of previous work that described how increased foliar damage from xanthomonad pathogens facilitated the growth of the human pathogen *Salmonella enterica* on tomato leaves (61). Improving the tolerance of food crops from lesion development as part of a multi-layered disease management strategy may help reduce yield losses and even prevent the colonization of human pathogens on diseased crops.

Although Xg and XgΔAvrHah1 do not differ measurably in *in planta* growth assays, it is clear that leaves infected with Xg "look" diseased while leaves infected with XgΔAvrHah1 do not, or have weaker, delayed symptoms. Given the choice, a plant without symptoms is obviously preferable. If the "tolerance" of the plant could be improved to prevent the formation of lesions, the damage from the pathogen may be reduced. Improving the tolerance of food crops from lesion development as part of a multi-layered disease management strategy may help reduce yield losses and prevent the colonization of human pathogens on diseased crops. In the case of AvrHah1-induced disease lesions, it would be possible to remove the EBEs from the promoters of genes that contribute to water soaking using new genome editing technology. First, the b*ona fide* S genes involved in AvrHah1-induced water soaking must be identified. The search for the targets of AvrHah1 that contribute to water soaking will be discussed in the following section.

5. Direct and indirect targets of AvrHah1 identified using RNA-seq complement water soaking in XgΔAvrHah1.

Background

Along with research conducted on TAL effectors from xanthomonad rice pathogens, much of the seminal research describing the functional activity of TAL effectors was done by Ulla Bonas and colleagues on the TAL effector AvrBs3. AvrBs3 induces cell hypertrophy on pepper, dependent on the AvrBs3 NLS, AD, and DBD, which likely plays a role in bacterial egress to leaf the surface (62). Indeed, loss of AvrBs3 was demonstrated to have a fitness cost for Xe in the field (40). AvrBs3 was originally found to induce the transcription of 11 plant genes, which were termed UPA genes (Upregulated by AvrBs3), nine of which were not activated in the presence of cycloheximide, a eukaryotic translation inhibitor (62). Activation of two UPA genes (UPA10 and UPA11) in the presence of cycloheximide indicated that AvrBs3 was likely their direct transcriptional activator because *de novo* translation was not required. A second search for UPA genes in pepper revealed that UPA20, a basic Helix Loop Helix (bHLH) transcription factor, was capable of inducing hypertrophy when transiently expressed in *N. benthamiana* (39). Furthermore, transient expression of AvrBs3 could not induce hypertrophy in *N. benthamiana* silenced for *upa20*, indicating that UPA20 was the *bona fide* S gene of AvrBs3. Because UPA20 is a transcription factor, the downstream targets of UPA20 are implicated in cell hypertrophy. UPA20 was found to activate expression to UPA7, an α-expansin, but the contribution of UPA7 to cell hypertrophy was not described (39).

Research on AvrBs3 targets led to the discovery of TAL effector Effector Binding Elements. Comparing the promoters of the known AvrBs3 direct targets (UPA10, UPA20, and the Resistance gene *Bs3*) revealed a common sequence motif termed the UPA Box, which partially encompassed the TATA box (39, 56). The identification of additional direct targets of AvrBs3 allowed for a more robust consensus sequence for the UPA Box (63). The discovery of the UPA box and the result that the central repeats of the DBD were required for TAL effector promoter binding in Electromobility Shift Assays (EMSAs) set the stage for the realization that the repeats, and likely the variable 12^{th} and 13^{th} amino acids in each repeat, were involved in DNA binding specificity (39).

The discovery of the nucleotide binding "code" for a given RVD was a major advancement in the field of TAL effector biology and for the development of TAL effector nucleases (TALENs), an important advancement in genome editing tools. The "uncoding" of the DNA binding capabilities of TAL effector RVDs was based on examples of TAL effectors and their known DNA binding sites (29, 30). For example, a repeat with an RVD "HD" will bind a cytosine, "NG" a thymine, "NI" an adenine with high probability. RVDs like "NN" are less stringent and bind a guanine or adenine. Some RVDs are more commonly observed than others, so data on their targeted nucleotide is not as robust.

Importantly, if the amino acid sequence of the TAL effector is know, the composite RVDs can be used to predict the identities of target EBEs in the host. In the case of AvrBs3, the discovery of the cDNA transcripts of UPA genes was done using suppressive subtractive

hybridization and northern analysis (39, 62). Current next-generation sequencing technologies have allowed for the use of RNA-sequencing (RNA-seq) to discover differentially expressed RNA transcripts. Thanks to computational tools that predict the presence of EBEs for a particular combination of RVDs (64), it is possible to scan an entire genome or promoter-ome for EBEs. These techniques partially depend on the availability of a high quality, well-annotated reference genome. Additionally, obtaining the plant germplasm that matches the reference genome cultivar is also ideal for RNA-seq analysis. For these reasons, we decided to use tomato Heinz 1706 as our host plant for an RNA-seq-based approach to find host gene targets of AvrHah1.

If a plant gene targeted by the TAL effector promotes pathogen growth or spread, the gene is considered a Susceptibility (S) gene. Identifying and characterizing S genes reveals pathogen strategies and is useful in the design of disease resistant plants, for example through the removal of relevant EBEs via DNA editing technologies (55). For TAL effectors that activate multiple host gene targets, such as those with EBEs that partially span a TATA box, it becomes increasingly challenging to identify the *bona fide* S gene(s) (31, 33). To probe single genes for pathogenicity functions, designer TAL effectors, or dTALEs, can be constructed and tested *in planta* for virulence contributions (32).

Plants have evolved diverse resistance mechanisms in response to TAL effectors (55). Some plants have strategically placed "EBE traps" in the promoters of resistance genes, as in the case of the pepper *Bs3* Resistance gene, which is activated at partially overlapping EBEs by the TAL effectors AvrBs3 and AvrHah1 (27, 56). Plants have accumulated mutations in promoter EBE regions that prevent successful activation of gene targets by TAL effectors, as in the case for rice *Os8N3* (65). Tomato plants utilize Bs4, a NB-LRR Resistance (R) Protein, to induce a cell death, or Hypersensitive Response (HR), in the presence of certain TAL effectors (66, 67).

In this section, we present the results of an RNA-seq analysis of tomato leaves infected with Xg wild type or XgΔAvrHah1 in order to identify candidate S genes of AvrHah1. Because many genes with EBEs were upregulated in the presence of AvrHah1, we used semi-quantitative RT-PCR to confirm AvrHah1-specific activation. We discovered that two basic Helix Loop Helix (bHLH) transcription factors were highly upregulated in the presence of AvrHah1. Using a transient luciferase reporter assay, we demonstrated that the promoters of the bHLH genes were activated by AvrHah1. We expanded our search to look for downstream targets of the transcription factors by mining our RNA-seq data for genes that were highly upregulated in the presence of AvrHah1 but without predicted EBEs. We showed that two pectin modifications genes—a pectate lyase and a pectinesterase—are downstream targets of both bHLHs using transient luciferase reporter assays. We constructed and used designer TAL effectors (dTALEs) that activate expression of the bHLHs. We showed that the bHLH dTALEs activated the expression of the pectate lyase and pectinesterase. We used the quantitative water soaking assay with dTALEs to determine the contribution of the bHLHs, pectate lyase, and pectinesterase in water soaking. We found that delivery of the bHLH and pectate lyase dTALEs by XgΔAvrHah1 complemented water soaking. Thus, the bHLHs are direct S genes and the pectate lyase is an indirect S gene. To our knowledge, the pectate lyase is the first example of an indirect target of a TAL effector shown to contribute to TAL effector symptoms.

Results

RNA-seq compares Xg wt- and XgΔAvrHah1-infected tomato revealing differentially expressed genes

We chose to do our RNA-seq analysis in tomato due to the availability of a high quality annotated tomato genome (68). All experiments were conducted in the sequenced tomato cultivar Heinz 1706 (UC Davis Tomato Genetics Resource Center). Our experimental strategy was to compare the gene expression profiles of tomato leaves infected with Xg wt or XgΔAvrHah1 and identify differentially expressed genes. We hypothesized that genes highly upregulated in Xg wt-infected tomatoes, but not XgΔAvrHah1-infected tomatoes, would be potential targets of AvrHah1. Additionally, we could narrow down the list of potential gene candidates by only considering genes with predicted EBEs for AvrHah1 in their promoters (Fig. 5-1). First, we computationally predicted the AvrHah1 EBEs in the tomato "promoterome", which we defined as the set of sequences 300bp upstream of annotated genes in the Heinz 1706 genome (68). Using the TALE-NT 2.0 algorithm (64) with a cutoff of four times the best possible score (3.76) and RVDs for AvrHah1 (NN IG NI NI NI HD HD NG NN NI HD HD HD NG), we found 4,106 possible EBEs in our promoterome on both the plus and minus strands, and 2,354 EBEs from just the plus strands. The majority of the predicted EBEs on the plus strand had high/weak binding scores of greater than 3 times the best possible score (Fig. 5-2). Two gene promoters had low/strong scores of less than 2 times the best possible score, and 24 gene promoters had intermediate scores of less than 2.5 times the best possible score.

The experimental design of the RNA-seq experiment was as follows: leaves from twelve Heinz 1706 tomato plants were syringe inoculated with either Xg wt or XgΔAvrHah1 (OD_{600}=0.25, totaling six plants for each bacterial strain). One hundred miligrams of inoculated leaf tissue was collected from three of the Xg wt or XgΔAvrHah1 tomato plants at 24hpi and frozen in liquid nitrogen and stored at -80^{o}C. Tissue from the remaining three plants was collected at 48hpi. RNA was extracted from the frozen tissue and used for Illumina TruSeq v2 library construction (with index adapters) and pooled for sequencing on a single lane on the Illumina HiSeq2000. RNA-sequencing reads were mapped to the Heinz 1706 genome (68) with CLC Genomics Workbench software after trimming and quality check. The mean expression values (in RPKM) of the three biological replicates per strain/time point were calculated and used to find differentially expressed genes between Xg wt and XgΔAvrHah1 with the Baggerley's test ($p < 0.05$). Because we found statistically significant differential expression in thousands of genes, we made the assumption that the S gene of AvrHah1 would be among the most highly upregulated in wt Xg-infected vs XgΔAvrHah1-infected tomato. Genes with significant differential expression were sorted by either fold change or RPKM. The most highly differential expressed genes were further sorted for the presence of AvrHah1 EBEs. Of the two time points, greater differences in gene expression were found at 48hpi.

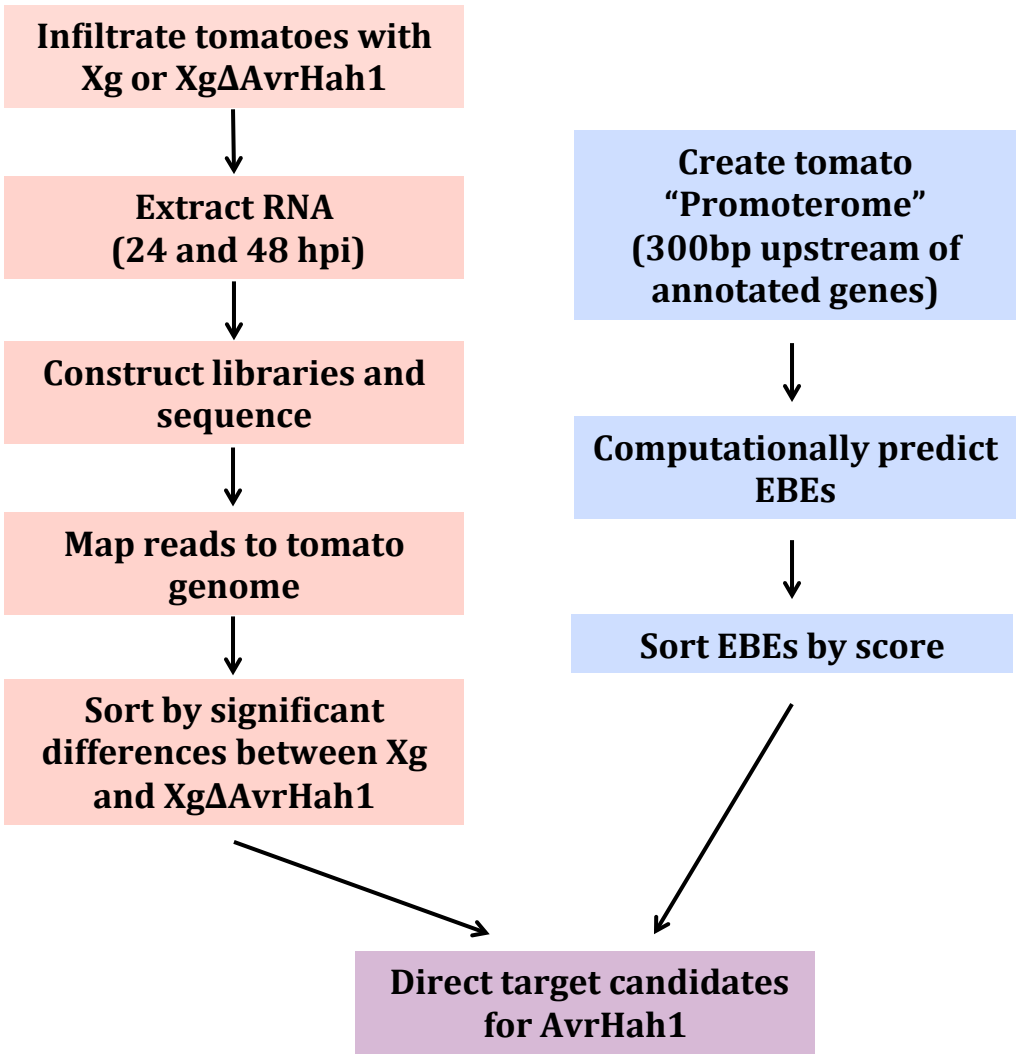


Figure 5-1. Design of RNA-seq experiment to find AvrHah1 direct target candidates
Tomatoes were syringe infiltrated at OD_{600} = 0.25. RNA-sequencing libraries were constructed using the Illumina TruSeq v2 library prep kit and libraries were sequenced on an Illumina HiSeq2000. RNA-sequencing analysis was done using the CLC Genomics Workbench.

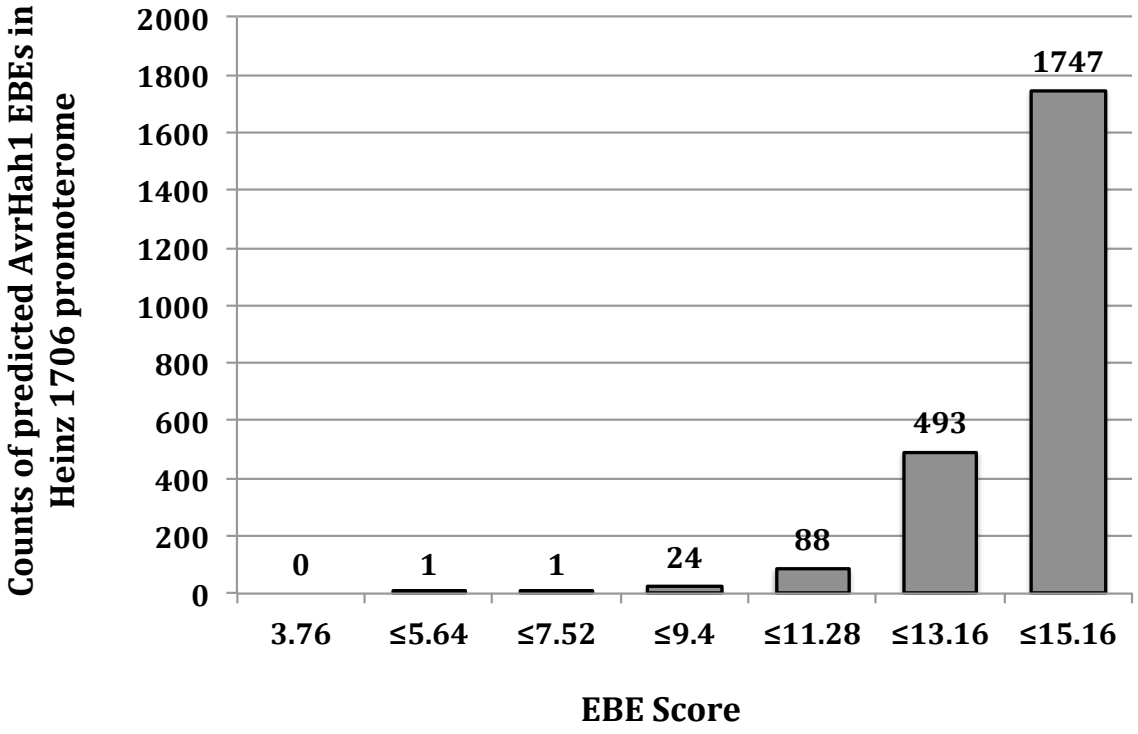


Figure 5-2. Predicted AvrHah1 EBEs in tomato promoterome organized by binding score

AvrHah1 predicted EBEs in the tomato promoterome (within 300bp of the start codon of annotated genes in Heinz 1706) using TALE-NT (64). Counts of EBEs with scores are indicated above the bars. The best possible score is 3.76.

Feature ID	EBE Score/bp from ATG	Fold Change	Mean RPKM		Predicted Protein Function
			Xg wt	XgΔAvrHah1	
Solyc11g069390		-2,089.68	202.31	0.1	Unknown
Solyc03g097820	3.96/108	-1,689.98	1,267.09	0.75	Transcription factor bHLH
Solyc06g072520	9.21/139	-1,515.11	568.68	0.38	Transcription factor bHLH
Solyc02g070180		-1,388.13	36.61	0.03	FAD-binding domain
Solyc02g089350		-1,279.44	997.81	0.78	Gibberellin regulated
Solyc03g116060		-1,092.12	155.13	0.14	Gibberellin-regulated
Solyc05g014000		-808.85	206.02	0.25	Pectate lyase
Solyc11g067180		-762.36	94.81	0.15	Fatty acyl coA reductase
Solyc02g084010		-714.68	99	0.14	SAUR-like
Solyc12g013850		-648.78	38.12	0.36	Glycosyl transferase
Solyc04g017720		-617.23	550.74	0.89	Gibberellin regulated
Solyc03g114430		-586.83	133.86	0.27	Unknown
Solyc02g070210	15.01/84	-572.75	36.11	0.06	Phosphatidylinositol transfer
Solyc07g006310		-533.28	129.78	0.24	Transcription factor
Solyc03g080100		-515.28	26.23	0.05	Copper P-type ATPase
Solyc12g055970		-428.36	6.6	0.02	Endoglucanase 1
Solyc11g019910		-398.78	89.27	0.22	Pectinesterase
Solyc03g033590		-391.4	55.85	0.14	SAUR-like protein
Solyc06g067910		-343.6	95.26	0.28	Os01g0611000-like protein
Solyc04g081870		-320.28	442.81	1.38	Expansin
Solyc04g016480		-316.36	34.33	0.11	Calmodulin binding

Table 7. Most highly differentially expressed genes from RNA-seq analysis of tomato infected with Xg wt- vs. XgΔAvrHah1

Genes are organized by fold change; negative values are because XgΔAvrHah1 was compared to Xg wt during analysis. Fold change values are displayed with heat map colors, where red and green indicate stronger and weaker differences, respectively. For genes that contain a promoter EBE, the binding score and distance from the ATG are listed.

Feature ID	Fold Change	Mean RPKM		EBE Score	EBE bp from ATG	Predicted Protein Function
		Xg	*XgavrHah1*			
Solyc03g097820	-1689.98	1267.09	0.75	3.96	108	Transcription factor bHLH
Solyc06g072520	-1515.11	568.68	0.38	9.21	139	Transcription factor bHLH
Solyc02g070210	-572.75	36.11	0.06	15.01	84	Phosphatidylinositol transfer
Solyc04g079700	-248.47	16.26	0.07	13.26	173	WD-40 repeat family protein
Solyc09g083100	-113.46	240.68	2.12	8.61	110	CIPK20
Solyc08g062450	-175.09	14.7	0.14	8.2	105	class II heat shock protein
Solyc06g066770	-93.43	134.46	1.44	10.33	109	Kelch repeat- F-box
Solyc06g049050	-56.52	10.03	0.18	14.9	137	Expansin
Solyc09g083090	-52.22	64.19	1.23	10.93	24	CIPK20
Solyc01g110440	-41.12	827.92	21.57	13.77	178	Arginine decarboxylase
Solyc05g053180	-20.89	30.21	1.45	14.98	272	Unknown
Solyc09g008650	-17.3	116.25	6.72	14.45	189	Peptidyl-prolyl cis-trans isomerase
Solyc09g008640	-31.2	29.09	1.78	13.37	262	Receptor-like protein kinase
Solyc02g082120	-9.91	35.9	3.62	11.55	82	Methyladenine glycosylase
Solyc06g071780	-8.78	19.15	2.18	11.87	100	DEHYDRATION-INDUCED19 5
Solyc11g007920	-9.01	618.16	70.34	6.16	75	Histone H2B
Solyc08g068790	-9.06	59.52	6.84	13.9	117	Tyramine hydroxycinnamoyl trans.
Solyc09g008240	-7.66	45.87	5.99	14.94	105	ABC transporter like
Solyc06g082440	-9.86	35.88	4.98	14.4	289	CIPK16
Solyc03g006700	-11.57	131.15	19.38	14.98	197	Peroxidase

Table 8. Differentially expressed genes with promoter EBEs from RNA-seq are direct target candidates for AvrHah1

Genes with both differential expression and predicted EBEs are direct target candidates of AvrHah1. Negative values in the fold change are because XgΔAvrHah1 was compared to Xg wt during statistical analysis. Fold change values are displayed with heat map colors, where red and green indicate stronger and weaker differences, respectively. EBE binding scores are displayed with heat map colors, where red and blue indicate stronger and weaker scores, respectively.

The majority of genes showing significant differential expression between Xg- and XgΔAvrHah1-infected tomato were found in both the 24hpi and 48hpi time points. Numerical values of gene expression discussed here will reflect the expression levels at 48hpi. The twenty most highly upregulated and differentially expressed genes are listed in Table 7. Of these twenty genes, three have predicted EBEs. The twenty most highly differentially expressed genes with promoter EBEs are listed in Table 8, and were therefore direct target candidates for AvrHah1. By collecting the EBEs from the top twenty differentially expressed genes in Table 8, the actual vs. predicted EBEs for AvrHah1 could be compared (Fig. 5-3). The observed EBEs show more variability than the predicted EBEs, particularly in the RVD/nucleotide pair HD/C that is commonly thought to have high specificity.

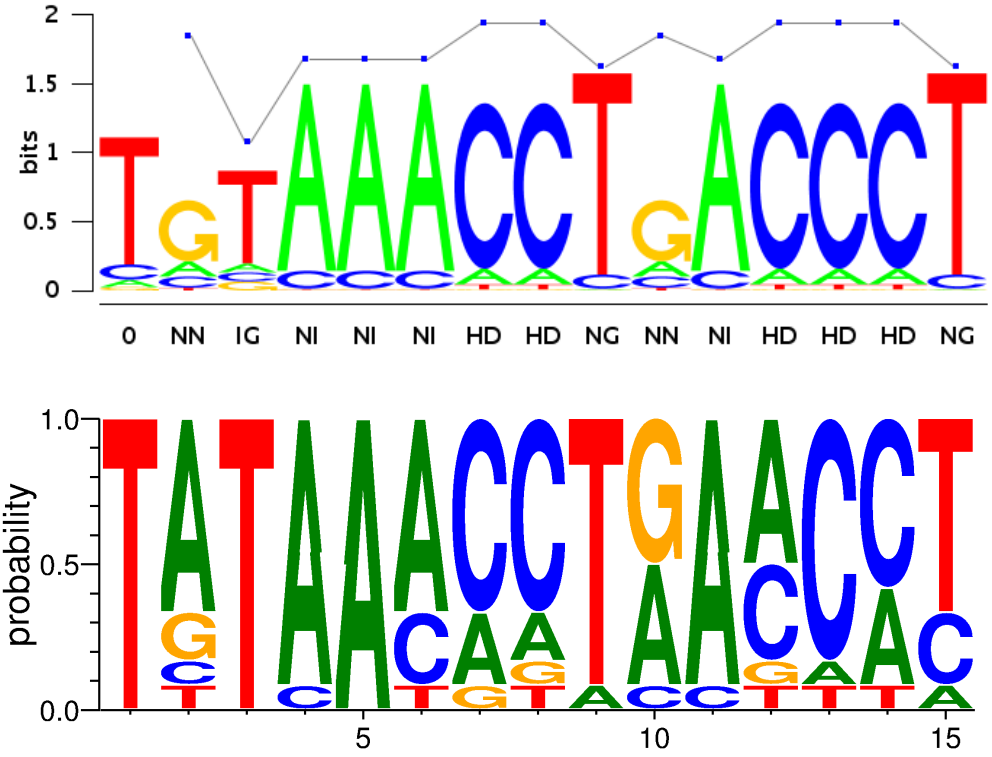


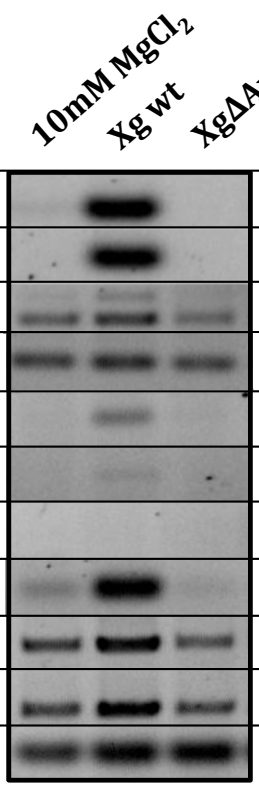
Figure 5-3. Predicted vs. observed AvrHah1 EBEs
Logo plots of EBEs were made using TALgetter (69). Top: predicted AvrHah1 EBE based on RVDs. Bottom: EBEs from the top 20 most highly differentially expressed genes identified in RNA-seq experiment.

Confirmation of AvrHah1 specific gene expression

Immediately apparent when comparing the highly upregulated genes in Table 7 with the direct target candidates in Table 8 are the two bHLH transcription factors, Solyc03g097820 and Solyc06g072520 that are both highly upregulated in Xg wt-infected tomato and possess EBEs with strong scores. We used semi-quantitative RT-PCR to confirm that activation of the bHLH genes and other putative AvrHah1 direct targets were specific for AvrHah1 (Fig. 5-4). We confirmed that expression of both bHLH genes was activated in tomato leaves infected with Xg, but not XgΔAvrHah1, at 24hpi. Additionally, the bHLHs were not active in mock infiltrated tissue (10mM $MgCl_2$).

Identification of two bHLH transcription factors as potential AvrHah1 direct targets

AvrBs3 upregulates a bHLH transcription factor in pepper (UPA20) that causes cell hypertrophy (39). Because we observed activation of a large number of genes in the presence of AvrHah1 in our RNA-seq experiment, we hypothesized that the two bHLH transcription factors were the S-gene targets of AvrHah1 and that they activated large transcriptional networks. To confirm that AvrHah1 has the ability to activate gene expression from the promoters of the bHLHs, we used an *Agrobacterium* luciferase reporter assay in *Nicotiana benthamiana*. We cloned the promoters of the bHLH transcription factors upstream of a luciferase reporter in a binary vector. *Agrobacterium* delivering the promoter:luciferase reporter was co-filtrated with *Agrobacterium* carrying a binary expression vector with AvrHah1, AvrBs3, or TAL20. TAL20 is a TAL effector from a xanthononad pathogen of cassava that activates a sucrose transporter (37). At 24hpi, 100μM luciferin was infiltrated into leaves, and six leaf discs per combination were collected and analyzed for luciferase activity. Although AvrHah1 and AvrBs3 are structurally different, they share some targets as determined by a partial overlap of their predicted EBE specificities. One of their shared targets is the Bs3 promoter (27). We used a Bs3 promoter:luciferase reporter as a positive control for activation by AvrHah1 and AvrBs3. For a negative control, we replaced the AvrHah1/AvrBs3 EBE in the Bs3 promoter for one that would match the RVD specificities of TAL20, creating Bs3Tal20 (37). As shown in Figure 5-4, the Bs3 promoter and the promoters of the bHLH genes (proSolyc03g097820 and proSolyc06g072520) are activated by both AvrBs3 and AvrHah1. Neither AvrHah1 nor AvrBs3 activates Bs3TAL20. TAL20 activates Bs3TAL20, but not the promoters of Bs3 or the promoters of the bHLHs. These results indicate that AvrHah1 has the ability to activate the promoters of the bHLH transcription factors. Therefore, the bHLH transcription factors are direct targets of AvrHah1.



	Function	EBE Score	bp from ATG	Fold change 24hpi	Fold change 48hpi
Solyc03g097820	TF bHLH	3.96	108	-277	-1,647
Solyc06g072520	TF bHLH	9.21	139	-1,288	-1,476
Solyc10g084430	Zinc finger	9.34	66	-86	-133
Solyc06g049050	Expansin	14.9	137	-31.09	-55.08
Solyc08g062450	Hsp20	13.74	220	-39	-193
Solyc02g070210	Pl transfer	15.01	84	-39.62	-557.2
Solyc08g077910	Expansin	13.98	58	-ns	-77
Solyc04g053000	SAUR protein	13.65	171	-28	-ns
Solyc09g083100	CBL kinase	8.61	110	-11	-110
Solyc06g066770	F-box protein	10.33	109	-44	-90

Figure 5-4. Semi-quantitative RT-PCR to confirm AvrHah1-specific activation of putative direct targets

Tomato leaves were infiltrated with 10mM $MgCl_2$ (negative control), Xg wt, or XgΔAvrHah1 (at OD_{600} = 0.25) and collected tissue for RNA extraction at 24hpi. Expression data from the RNA-seq experiment is also listed. (ns = not significantly different during that time point).

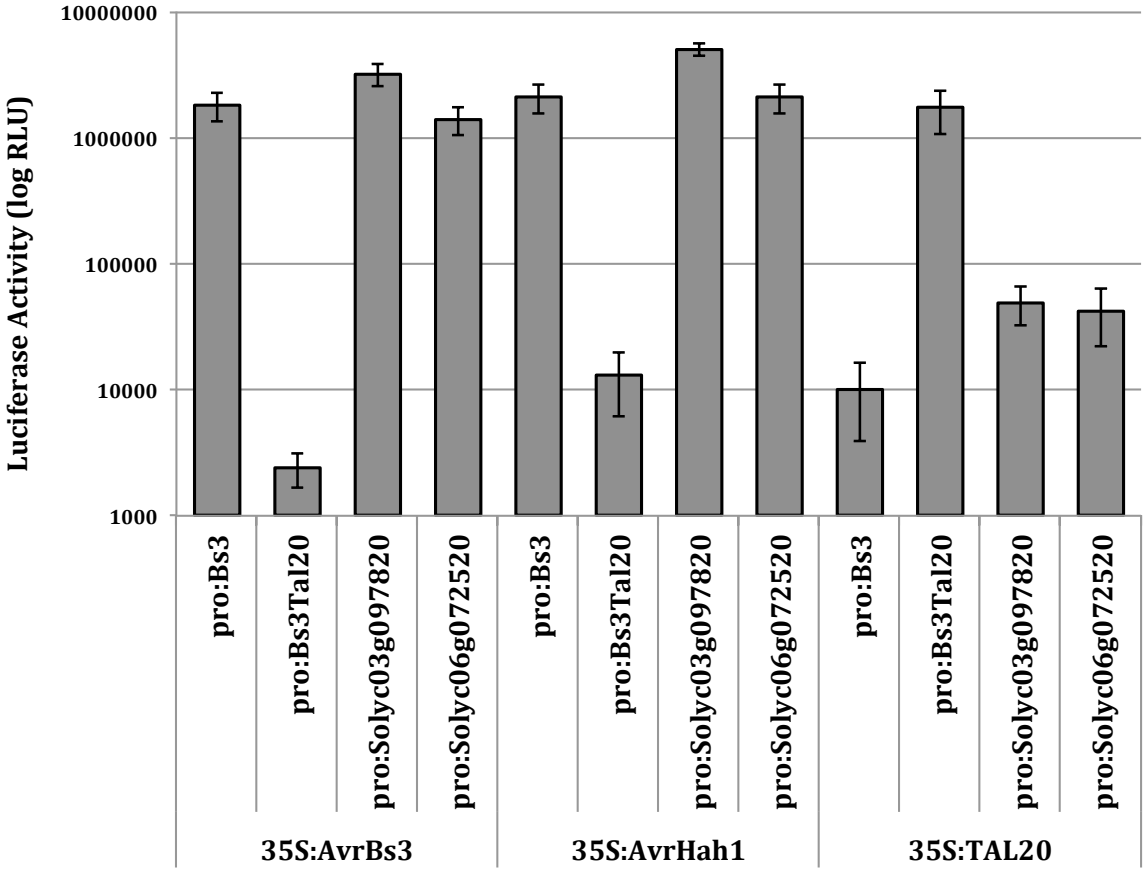


Figure 5-5. bHLH transcription factor promoters are activated by AvrHah1 in transient luciferase reporter assays

N. benthamiana was co-infiltrated with separate *Agrobacterium* strains carrying a promoter:luciferase reporter (in pGWB35) and a TAL effector (in p1776) (OD_{600} = 0.4 for each strain). At 24hpi the leaves were infiltrated with 100µM luciferin and six leaf discs were collected for luciferase activity using a Wallace Envision Plate Reader.

Identification of indirect targets of AvrHah1

We hypothesized that genes upregulated by the bHLH transcription factors would be among the genes we identified in our RNA-seq experiment as differentially upregulated but without AvrHah1 EBEs. Targets of the bHLHs could therefore be considered "indirect" targets of AvrHah1. We did semi-quantitative RT-PCR on 48hpi infected tomato leaves to check for the expression of genes without EBEs but present in our list of highly upregulated genes (Table 7). We found that two tomato genes involved in pectin modification, a pectate lyase (Solyc05g014000) and a pectinesterase (Solyc11g019910), were upregulated in the presence of Xg wt but not XgΔAvrHah1 (Fig. 5-6). Activation of these genes was restored in the XgΔAvrHah1 + AvrHah1 complementation strain.

To determine if the promoters of the pectin modification genes could be activated by the bHLH transcription factors, we cloned the promoters of the pectate lyase and pectinesterase for analysis in the luciferase reporter assay (Fig. 5-7). We co-infiltrated these luciferase reporters with a binary expression vector containing AvrHah1 or a cDNA transcript of either bHLH transcription factor. We observed that luciferase activity was significantly increased above background levels when the promoters of the pectate lyase and the pectinesterase were co-expressed with either bHLH transcription factor.

Interestingly, we also observed increased luciferase activity from the pectate lyase and pectinesterase promoters in response to AvrHah1 in *Nicotiana benthamiana*. These results indicate that AvrHah1 may be activating the expression of endogenous bHLH transcription factors in *N. benthamiana* that are also capable of activating the pectin modification promoters. Given that AvrHah1 causes water soaking in pepper, tomato, and *N. benthamiana*, it is likely that the S gene targets that contribute to water soaking are homologous between these three host plants. Indeed, we identified two bHLH transcription factors in *N. benthamiana* that have predicted AvrHah1 EBEs with strong scores (3.96 out of best possible 3.76, Niben101Scf00376g01004.1, 162bp upstream of the ATG and Niben101Scf01182g03011.1, 170bp upstream of the ATG).

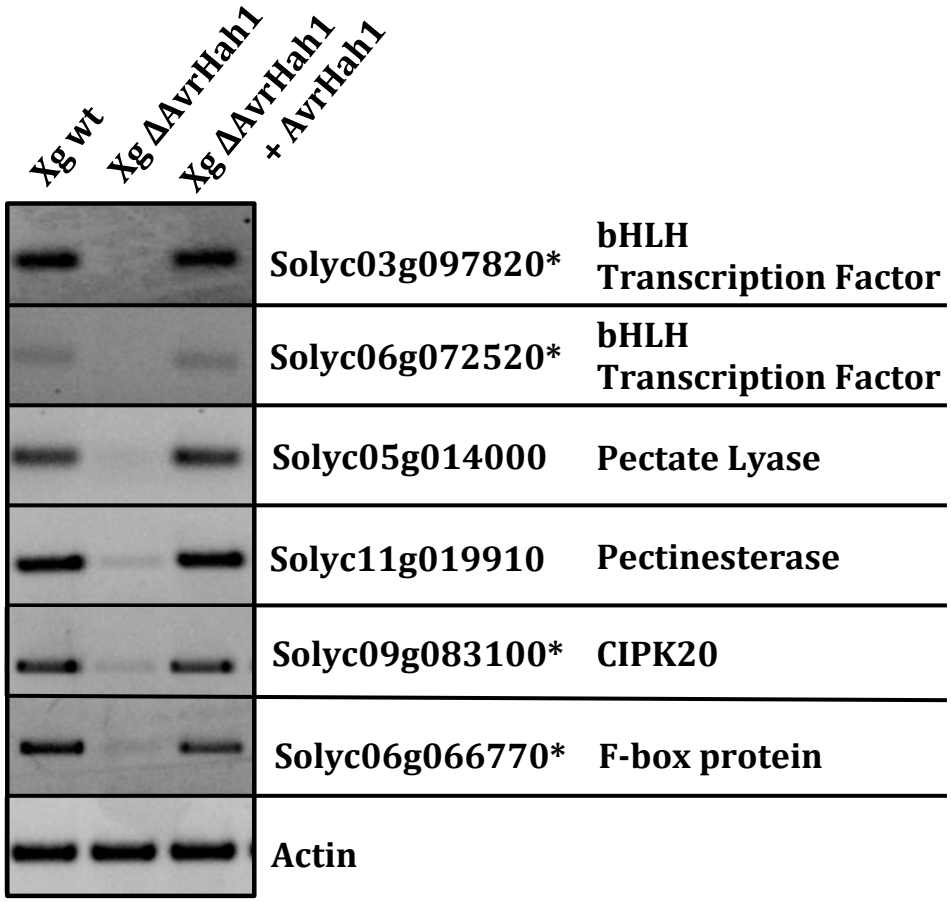


Figure 5-6. Two pectin modification genes without EBEs are upregulated in the presence of AvrHah1

Semi-quantitative RT-PCR for genes listed on the right. Tomato leaves were infiltrated with Xg wt, XgΔAvrHah1, or XgΔAvrHah1 + AvrHah1 at OD_{600} = 0.25. Tissue was collected for RNA extraction at 48hpi and tissue was collected for RNA extraction at 24hpi. *Indicates a predicted EBE for AvrHah1 in the gene promoter.

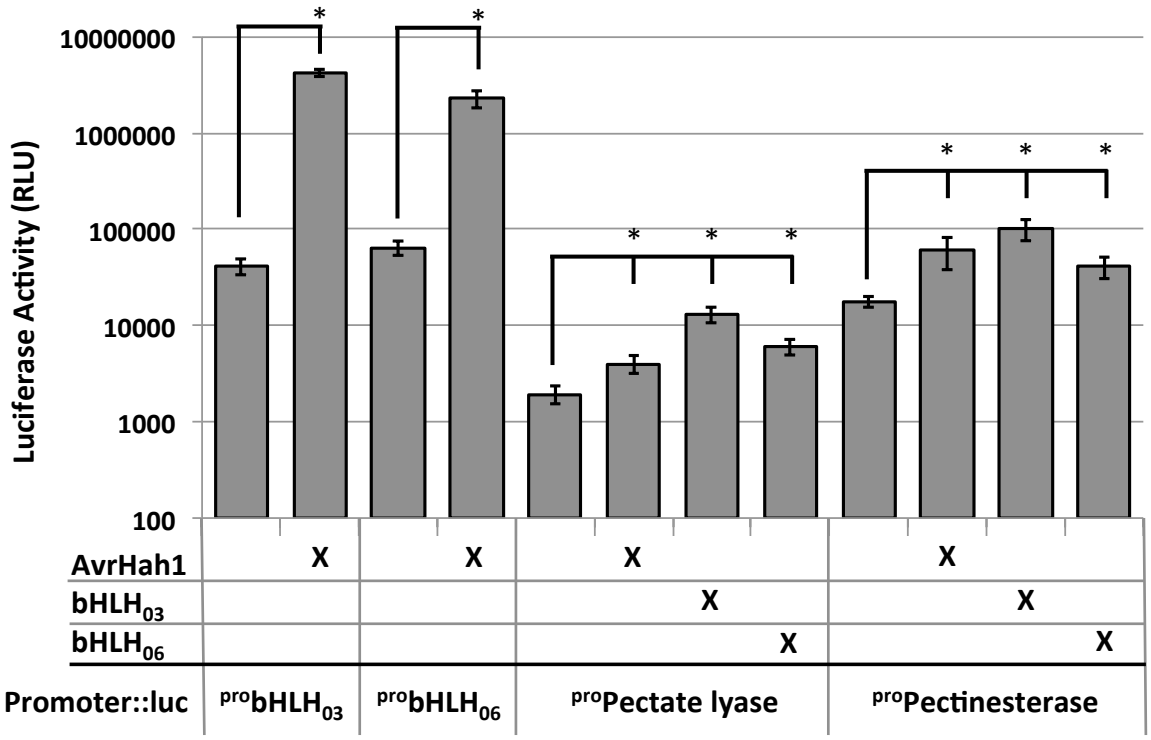


Figure 5-7. The pectin modification promoters are activated by AvrHah1 and the bHLH transcription factors in transient luciferase reporter assays
N. benthamiana was co-infiltrated with separate *Agrobacterium* strains carrying a promoter:luciferase reporter (in pGWB35) and a TAL effector or bHLH transcription factor (in p1776)) (OD_{600} = 0.4 for each strain). At 24hpi the leaves were infiltrated with 100µM luciferin and six leaf discs were collected for luciferase activity using a Wallace Envision Plate Reader. *Significant differences from promoter background were calculated with an unpaired student t-test ($p < 0.05$). pro = promoter, luc = luciferase, bHLH03 and bHLH06 are cDNA transcripts of Solyc03g097820 and Solyc06g072520, respectively.

Designer TAL effectors activate AvrHah1 targets

We constructed designer TAL effectors (dTALES) to activate the expression of the bHLH transcription factors and the pectin modification genes in order to determine their contributions to water soaking. dTALEs were constructed as previously described, using the RVDs NI, NG, HD, and NN to target the nucleotides A, T, C, and G, respectively (32). Information about the target EBEs in the tomato promoters and the RVDs used to target the EBEs can be found in Table 9. All dTALEs are driven by the TAL20 promoter (37) in the broad host-range vector pVSP61. Plasmids containing dTALEs were conjugated into XgΔAvrHah1. Two dTALEs per target gene were constructed. The bHLH transcription factor Solyc03g097820 is targeted by dT 504 and dT 505. The bHLH transcription factor Solyc06g072520 is targteted by dT 505 and dT 506. The pectate lyase Solyc05g014000 is targeted by dT 512 and dT 513. The pectinesterase Solyc11g019910 is targeted by dT 514 and dT 515.

We used semi-quantitative RT-PCR to check for activation of dTALE gene targets when delivered by XgΔAvrHah1 (Fig. 5-8). As expected, both bHLH transcription factors and both pectin modification genes are expressed in tomato infected with XgΔAvrHah1 + AvrHah1. No expression is observed with XgΔAvrHah1 + AvrHah1$_{\Delta DBD}$ or XgΔAvrHah1 + AvrBs3 (except actin). Expression of the bHLH transcription factor Solyc03g097820 was activated in response to dT 504 and dT 505. Expression of the bHLH transcription factor Solyc06g072520 was activated in response to dT 506 and dT 507. Importantly, gene expression of the pectate lyase Solyc05g014000 was observed in response to the two dTALEs activating each bHLH transcription factor. These results provide further evidence that the pectate lyase is a downstream target of both bHLH transcription factors. Compared to expression of the pectate lyase, weak expression of the pectinesterase Solyc11g019910 was observed in response to the dTALEs for the bHLH transcription factors.

Only one of the two dTALEs constructed for the pectate lyase and pectinesterase was able to activate target gene expression: dT 513 for the pectate lyase and dT 514 for the pectinesterase. Weak activation of the pectate lyase was observed in response to dT 513 in comparison to the expression levels observed when activated by AvrHah1 and the bHLH transcription factor dTALEs. dT 514 activated the pectinesterase strongly compared to the dTALEs for the bHLH transcription factors. It is possible that the targeted EBEs for dT 512 and dT 515 were too distant from the start codon to activate gene expression (Table 9).

Because only one of the dTALEs for the pectate lyase (dT 513) and one for the pectinesterase (dT 514) activated gene expression of the desired target, we selected these dTALEs, along with a single dTALE for each bHLH transcription factor, dT 505 (Solyc03g097820) and dT 506 (Solyc06g072520), for further analysis (Fig. 5-9).

Tomato target	dTALE ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Score: Actual/Best	bp from ATG
Solyc03g097820 (bHLH Transcription Factor)	dT 504	A	T	C	T	C	T	C	T	C	T	T	C	A	T	C	T	T	T	4.04/4.04	129
		NI	NG	HD	NG	HD	NG	HD	NG	HD	NG	NG	HD	NI	NG	HD	NG	NG	NG		
	dT 505	A	T	A	G	A	T	A	T	A	A	G	C	T	A	C	C	A	G	5.45/5.45	274
		NI	NG	NI	NN	NI	NG	NI	NG	NI	NI	NN	HD	NG	NI	HD	HD	NI	NN		
Solyc06g072520 (bHLH Transcription Factor)	dT 506	A	T	A	C	A	G	G	A	T	A	T	C	C	C	T	T	T	C	4.96/4.96	166
		NI	NG	NI	HD	NI	NN	NN	NI	NG	NI	NG	HD	HD	HD	NG	NG	NG	HD		
	dT 507	A	T	C	A	T	C	A	A	G	C	T	T	C	T	C	C	C	A	4.16/4.16	237
		NI	NG	HD	NI	NG	HD	NI	NI	NN	HD	NG	NG	HD	NG	HD	HD	HD	NI		
Solyc05g014000 (Pectate Lyase)	dT 512	A	G	A	C	C	C	A	G	C	A	G	A	C	T	C	A	A	G	5.74/5.74	429
		NI	NN	NI	HD	HD	HD	NI	NN	HD	NI	NN	NI	HD	NG	HD	NI	NI	NN	14.02/5.74	592
	dT 513	A	C	T	A	A	A	A	G	T	A	T	T	C	A	C	A	T	C	15.23/5.74	189
		NI	HD	NG	NI	NI	NI	NI	NN	NG	NI	NG	NG	HD	NI	HD	NI	NG	HD	4.22/4.22	381
Solyc11g019910 (Pectinesterase)	dT 514	A	T	T	T	C	C	C	T	C	A	C	T	A	A	T	A	C	T	3.73/3.73	420
		NI	NG	NG	NG	HD	HD	HD	NG	HD	NI	HD	NG	NI	NI	NG	NI	HD	NG		
	dT 515	A	C	C	A	A	C	T	A	A	C	A	A	T	A	T	C	A	T	3.44/3.44	565
		NI	HD	HD	NI	NI	HD	NG	NI	NI	HD	NI	NI	NG	NI	NG	HD	NI	NG		

Table 9. Specifications of dTALEs

Best possible scores determined using TALE-NT (64).

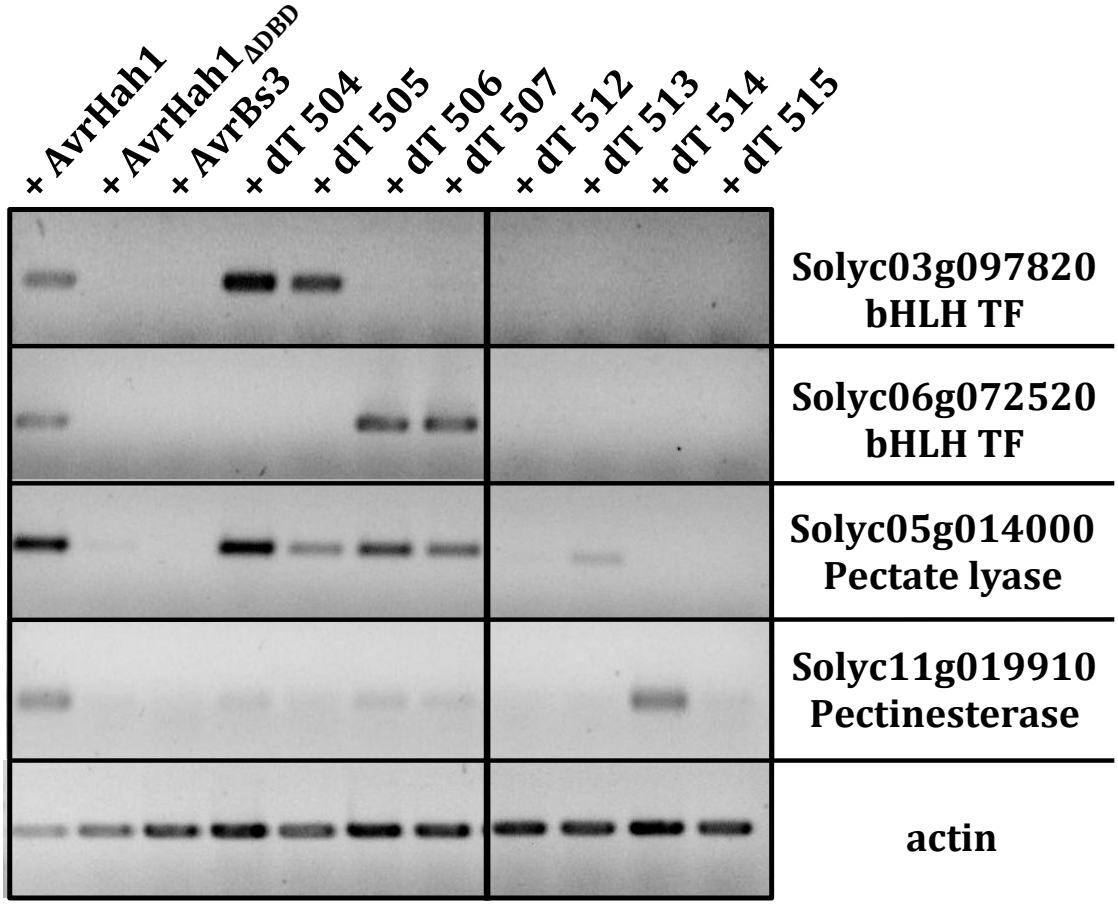


Figure 5-8. Confirmation of *in planta* gene activation by dTALEs
Semi-quantitative RT-PCR was performed on tomato Heinz 1706 infiltrated with XgΔAvrHah1 + dTALE strains (24hpi, OD_{600} = 0.25). Targeted promoters for direct gene activation are as follows: Solyc03g097820: dTs 504 and 505, Solyc06g072520: dTs 506 and 507, Solyc05g014000: dTs 512 and 513, Solyc11g019910: dTs 514 and 515.

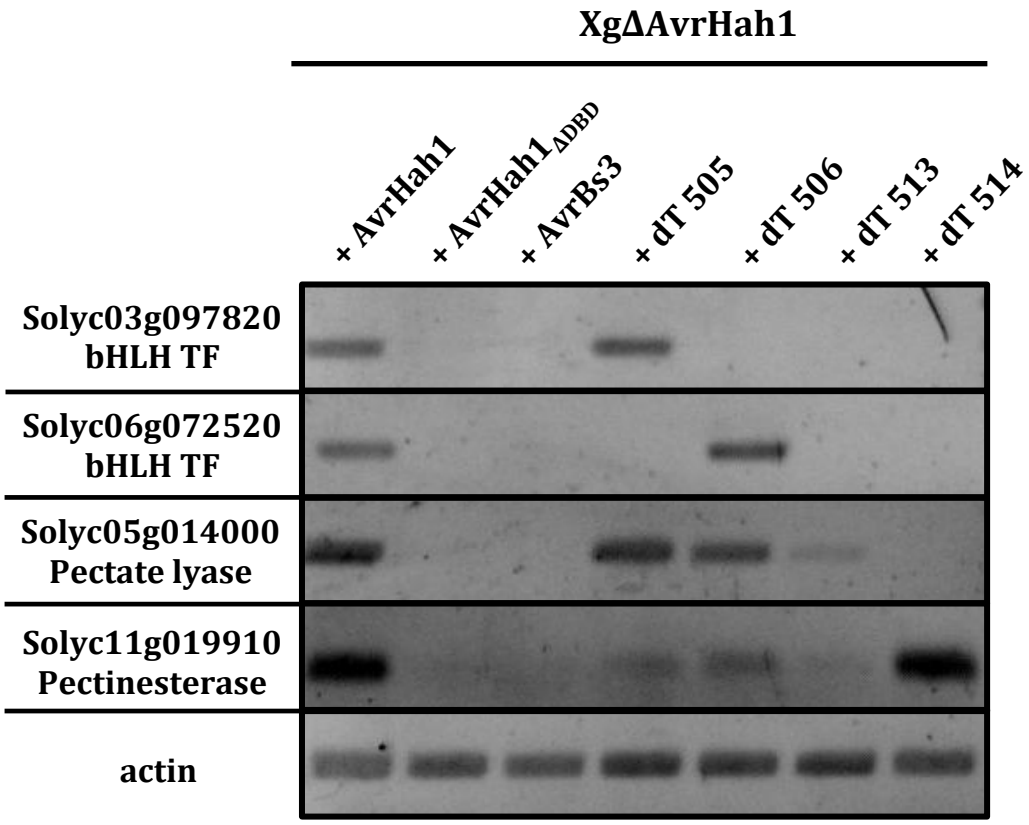


Figure 5-9. Summary of dTALE gene expression activities
Semi-quantitative RT-PCR was performed on tomato Heinz 1706 infiltrated with XgΔAvrHah1 + dTALE strains at OD_{600} = 0.25. Leaf tissue for RNA-extraction was collected at 24hpi. Targeted promoters for direct gene activation are as follows: Solyc03g097820: dT 505, Solyc06g072520: dT 506, Solyc05g014000: dT 513, Solyc11g019910: dT 514.

Quantitative water soaking assay using dTALEs for AvrHah1 targets

To determine if the targets of the dTALEs contribute to water soaking, we performed a quantitative water soaking assay (similar to Fig. 4-10) with strains of XgΔAvrHah1 + dTALEs (Fig. 5-10). Leaves were syringe infiltrated with Xg strains at OD_{600} = 0.1. At 48 hpi, leaves were submerged in water for 20 minutes to promote intake of water into the apoplast (no wounds were purposefully introduced). Prior to submerging leaves in water, infiltrated zones showed no noticeable phenotype, except for the HR in response to AvrBs3. The apoplastic fluid from two 0.5cm^2 leaf discs was collected by centrifugation and weighed. For comparison, we also sampled apoplastic fluid from 'untreated' tomato leaves (also submerged in water and blotted dry) and leaves completely syringe infiltrated with water immediately prior to collection.

Little apoplastic fluid is collected from untreated leaves. In contrast, leaves infiltrated with water experience an approximate 6-fold increase in apoplastic fluid (by weight). Leaves infiltrated with Xg wt experience an approximate 8-fold increase in apoplastic fluid compared to untreated leaves. Leaves infiltrated with XgΔAvrHah1 show a dramatic reduction in apoplastic fluid intake compared to Xg wt, however with a significant increase in apoplastic fluid above the level of untreated leaves. These results indicate that Xg may have other virulence factors that have a minor contribution to water soaking. Although the reduced water soaking of XgΔAvrHah1 was not fully complemented by AvrHah1 to wild type levels (XgΔAvrHah1 + AvrHah1), it was similar to water-infiltrated leaves. The zones infiltrated with XgΔAvrHah1 + AvrBs3 showed a weak cell death response and little apoplastic fluid (similar to untreated leaves).

dTALEs activating the bHLH transcription factors (dT 505 and dT 506) complemented water soaking to the level of XgΔAvrHah1 + AvrHah1. Interestingly, dT 513 activating the pectate lyase also complemented water soaking. Apoplastic fluid weights for the pectinesterase (dT 514) were indistinguishable from those of XgΔAvrHah1, indicating that the pectinesterase does not complement water soaking. A 1:1 combination of dT 505 and dT 506 (final OD_{600} = 0.1) showed similar water soaking to the individual dTALE levels. A mixture of dT 513 and dT 514 showed a water soaking level intermediate to its two individual dTALEs, consistent with the result that dT 513, but not dT 514, contributes to water soaking. These results provide evidence that the bHLH transcription factors are *bona fide* direct S gene targets of AvrHah1 as they are able to complement water soaking. Additionally, our experimental evidence points to the pectate lyase as an indirect S gene target of AvrHah1. To our knowledge, this is the first instance of a TAL effector-activated transcription factor target with demonstrated virulence activity.

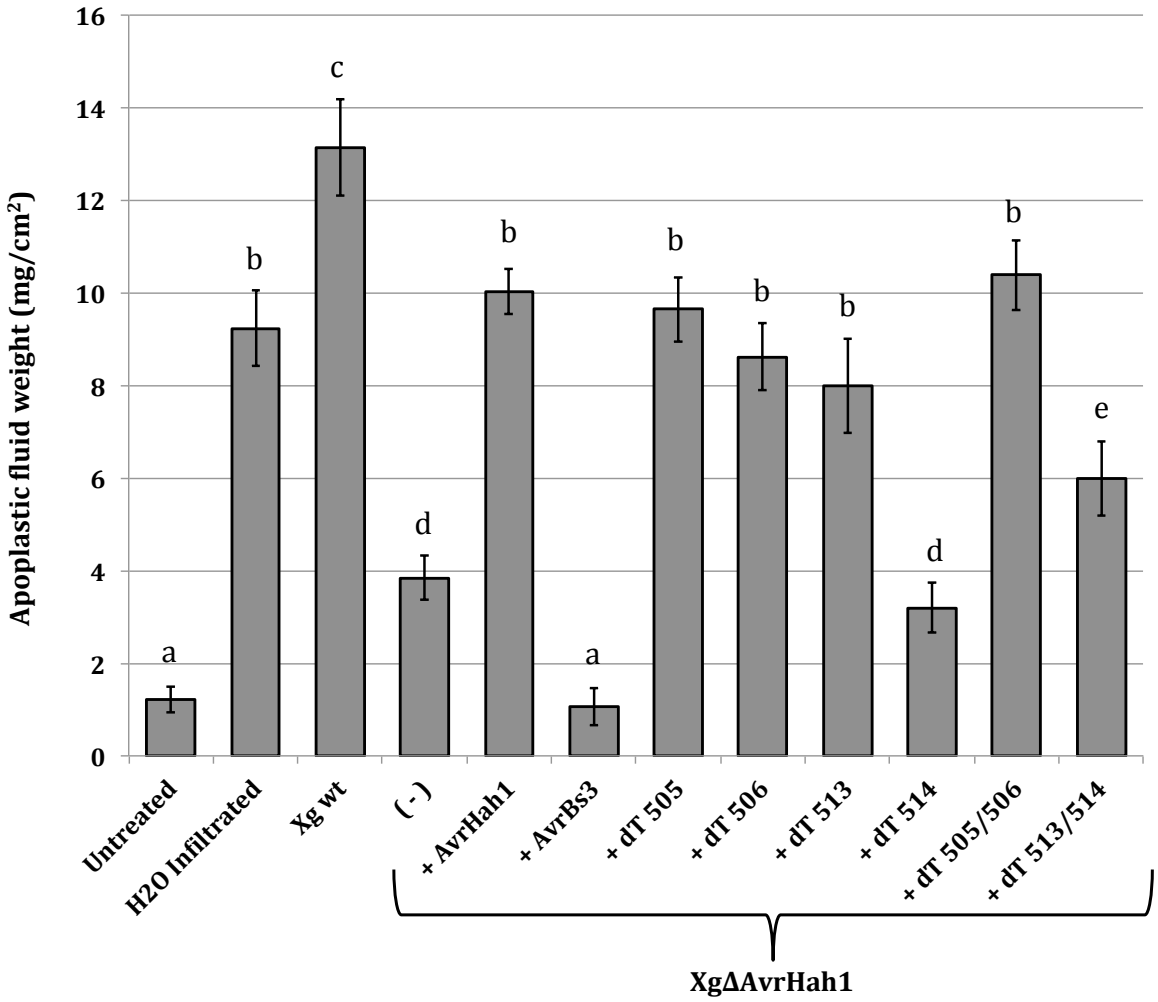


Figure 5-10. Contribution of direct and indirect AvrHah1 targets to water soaking
Tomato leaves were syringe infiltrated with a concentrated inoculum (OD_{600} = 0.1, ~10^8 CFU/mL). At 48hpi, leaves were submerged in water for 20 minutes and apoplastic fluid was collected and weighed. Average weights and standard errors from 12 samples (consisting of two 0.5cm^2 leaf discs) are shown. Significant differences (determined by unpaired student t test) are indicated by different letters ($p < 0.05$).

Discussion

To find the downstream targets of AvrHah1 responsible for water soaking, we used RNA-seq to compare the gene expression profiles of Xg and XgΔAvrHah1-infected tomato leaves. We identified two highly upregulated bHLH transcription factors, Solyc03g097820 and Solyc06g072520, that share 66% amino acid identity and are 86% and 64% similar to UPA20—the bHLH target of AvrBs3 in pepper (39). Solyc03g097820 and Solyc06g072520 are in a sister clade and part of the bHLH subfamily 1 (70). Solyc03g097820 (annotated as bHLH022) was found to be expressed preferentially in young fruits (70). Solyc06g072520 (annotated as bHLH048) has been found to be a drought responsive gene in a drought-tolerant tomato variety (71).

Because UPA20 induces the cell hypertrophy phenotype of AvrBs3, we were interested to test whether the bHLH transcription factors activated by AvrHah1 contribute to water soaking. Transient expression of both AvrBs3 and UPA20 in pepper induces cell hypertrophy in *Nicotiana benthamiana* (39), however we could not induce water soaking by transiently expressing AvrHah1 or the bHLH transcription factors in *Nicotiana benthamiana* (data not shown), suggesting that AvrHah1-induced water soaking needs to occur in the context of the pathogen. RT-PCR in pepper demonstrated that an α-expansin (UPA7) was upregulated when UPA20 was transiently expressed (39). UPA7 is therefore an indirect target of AvrBs3, but a role for UPA7 in cell hypertrophy was not reported. Here, we provide evidence that a downstream target of transcription factors activated by a TAL effector contributes to symptom development.

We selected genes in our RNA-seq dataset that were upregulated in the presence of AvrHah1 but without predicted AvrHah1 EBEs. After confirming activation using semi-quantitative RT-PCR, we selected two pectin modification genes for further study, a pectate lyase (Solyc05g014000) and a pectinesterase (Solyc11g019910). We were interested in pectin modification because several examples have demonstrated the importance of pectin in the interactions between plant pathogens and their hosts (72, 73). We hypothesized that cell wall modifications were mediating the absorption of water into the apoplast. We demonstrated that the pectate lyase and pectinesterase are activated by the bHLH transcription factors using a transient luciferase reporter assay and semi-quantitative RT-PCR with dTALEs for the bHLH transcription factors. Importantly, the dTALEs for both bHLH transcription factors and the pectate lyase were able to complement water soaking in XgΔAvrHah1 to the level of XgΔAvrHah1 + AvrHah1, indicating that they are *bona fide* S gene targets of AvrHah1.

The ability of the pectate lyase dTALE to complement water soaking points to the importance of cell wall modification in symptom development. One possibility is that by changing the composition of pectin, the hygroscopicity of the cell wall is increased, allowing for the fast absorption of water through natural wounds or cracks in the epidermis (Fig. 5-11). Future experiments studying changes in cell wall composition of Xg infected leaves will further identify the mechanism by which AvrHah1 and the pectate lyase cause water soaking.

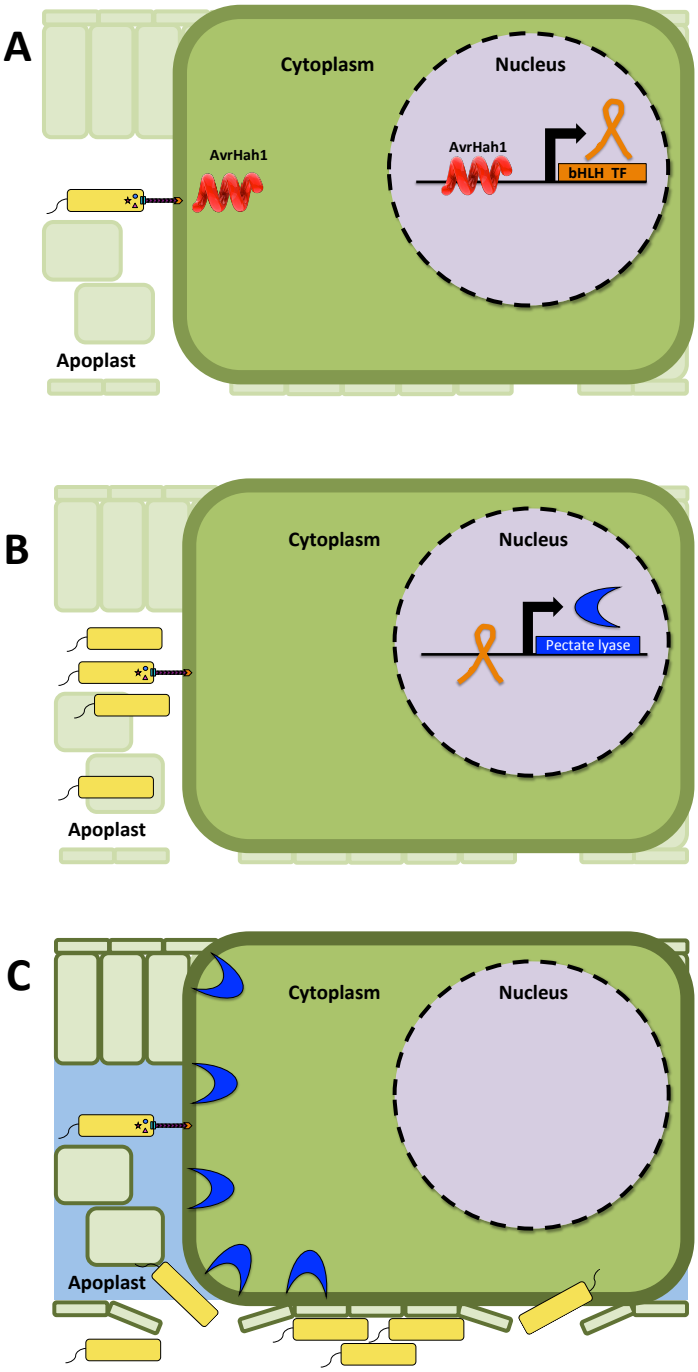


Figure 5-11. Model of AvrHah1-induced water soaking
(A) Xg secretes AvrHah1 into plant cells, which activates the expression of bHLH transcription factor genes (B) bHLH transcription factors activate the expression of the pectate lyase gene. Meanwhile, Xg continues to grow in the apoplast. (C) Pectate lyase protein activity modifies the cell wall such that it absorbs water from the outside of the leaf. Xg cells are able to escape from the apoplast in higher number through wounds or natural openings.

6. Differential recognition of AvrBs3 and AvrHah1 by the tomato R protein Bs4

Background

Bs4 is a TIR-NB-LRR type Resistance protein in tomato that recognizes the TAL effectors AvrBs4, Hax3, and Hax4 (66, 67). Previous work demonstrated that AvrBs3 was also recognized by Bs4, but only when AvrBs3 was delivered via *Agrobacterium* with a strong promoter (and not when delivered via *Xanthomonas*) (74). We were surprised to observe that in our system AvrBs3 induced a cell death in tomato when delivered by XgΔAvrHah1 or Xe85-10 (Figs. 4-9 and 4-6, respectively). Evidence points to a direct recognition model between Bs4 and its recognized TAL effectors, likely involving the repeats of the DBD (66). Interestingly, we did not observe a Bs4-dependent recognition of AvrHah1, indicating that there may be differential recognition between these two TAL effectors.

Results

Delivery of AvrBs3 by XgΔAvrHah1 induces a cell death, or Hypersensitive Response (HR), in tomato (Fig. 6-1). We hypothesized that Bs4 also recognizes AvrHah1, but the water soaking induced by AvrHah1 disrupts the cell death response. To disprove this hypothesis, we truncated the last 46 aa of AvrHah1 to delete the Activation Domain (AD), which removed the ability of AvrHah1 to cause water soaking but maintained the DBD. Neither delivery of AvrHah1$_{\Delta AD}$ nor AvrHah1$_{\Delta DBD}$ induced a cell death response in tomato (Fig 6-2), indicating that Bs4 differentially recognizes AvrHah1 and AvrBs3.

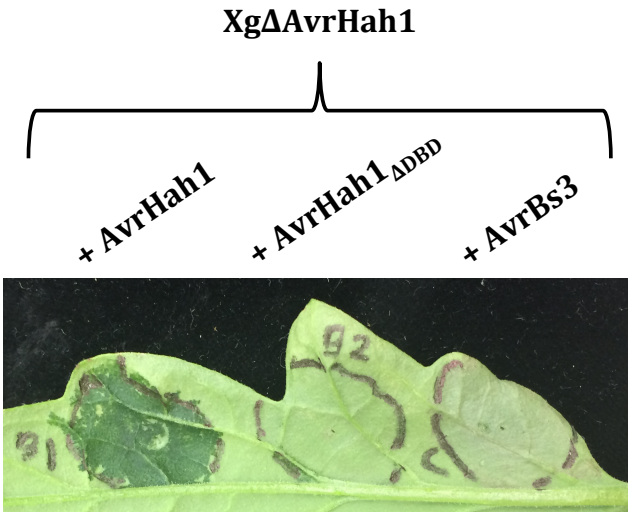


Figure 6-1. AvrBs3 elicits HR in tomato, whereas AvrHah1 elicits water soaking
XgΔAvrHah1 was infiltrated into tomato Heinz 1706 (OD_{600} = 0.1, 48 hpi) and submerged in water for 20 minutes for allow water soaking.

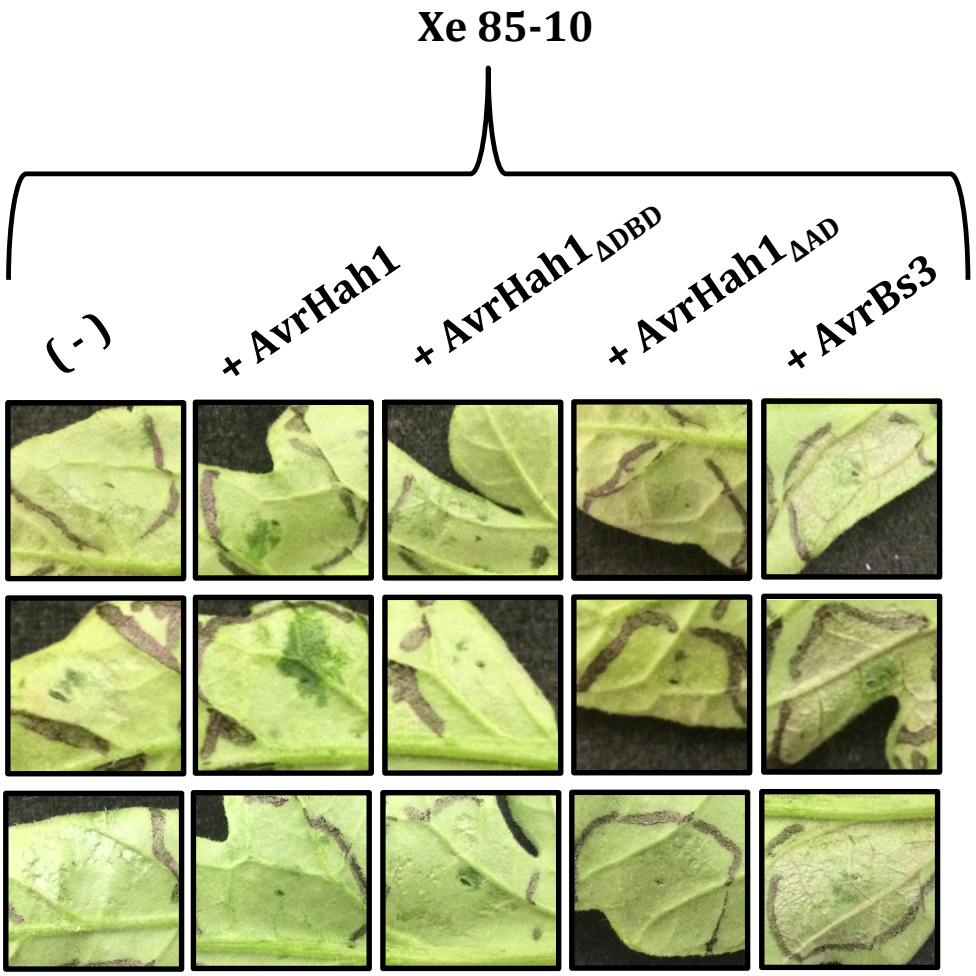


Figure 6-2. Two non-functional AvrHah1 mutants are unrecognized by Bs4
Xe85-10 carrying the TAL effector indicated was infiltrated into tomato Heinz 1706 (OD_{600} = 0.1, 48 hpi). Leaves were not submerged in water. DBD = DNA Binding Domain, AD = Acticvation Domain.

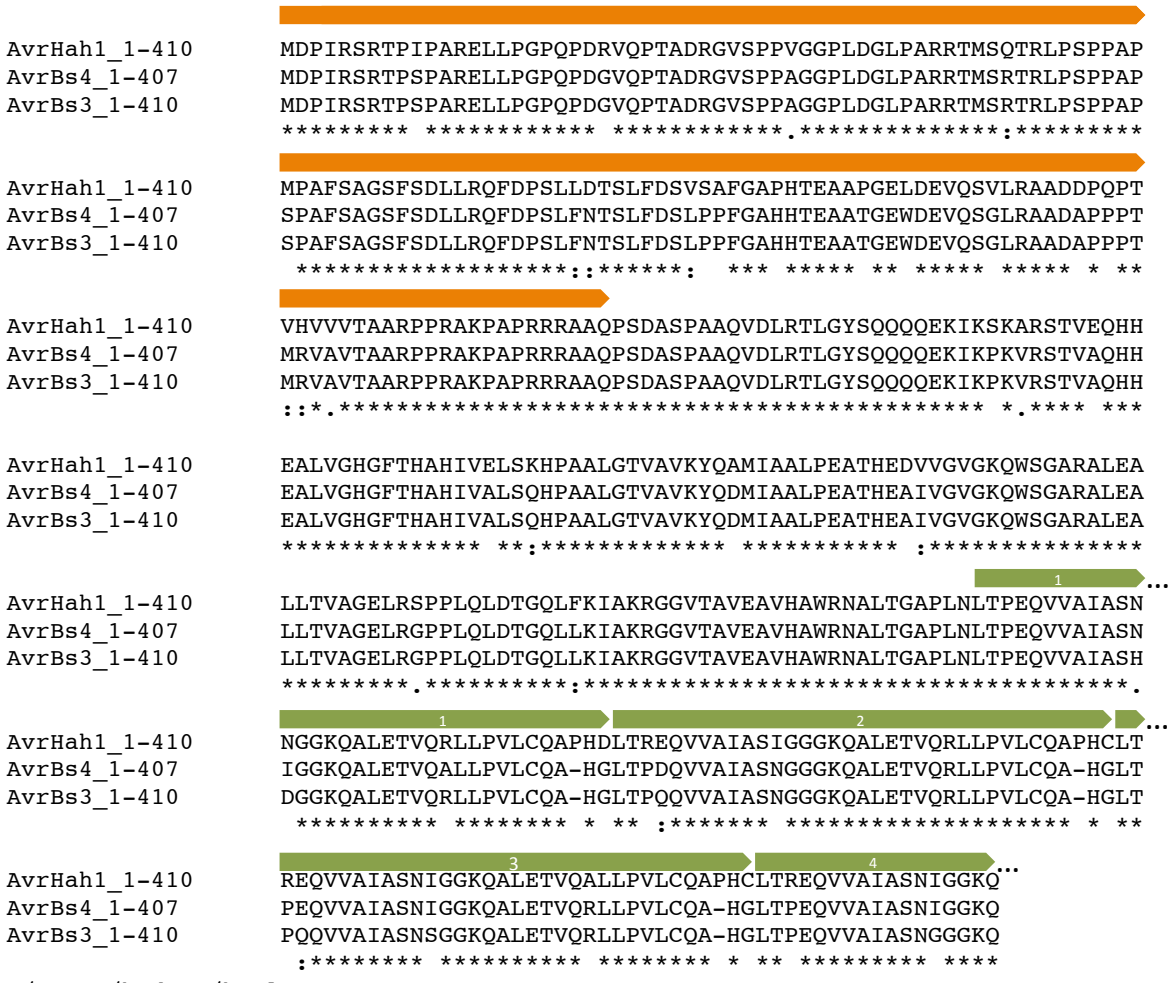


Figure 6-3. Amino acid alignment of the proposed Bs4 minimal recognition domain
A truncation of all but the first 3.5 repeats of AvrBs4 was sufficient to activate Bs4 recognition (66). An orange arrow depicts The Type III Secretion signal (T3SS) and green arrows depict the individual repeats of the DBD. The 35 aa repeats of AvrHah1 have an extra proline in position 33 and a different terminal amino acid (Aspartate or Cysteine) from the 34 aa repeats of AvrBs3 and AvrBs4 (Glycine).

Discussion

We observe a cell death response in tomato Heinz 1706 when AvrBs3, but not AvrHah1, is delivered by XgΔAvrHah1. We show that two non-functional mutants of AvrHah1 (lacking the DBD or the AD) remain unrecognized in tomato, indicating that any water soaking activity conferred by AvrHah1 is not disrupting a Bs4-mediated HR. These two TALEs were under the control of the same promoter, suggesting that Bs4 recognizes AvrBs3 but not AvrHah1.

An amino acid alignment of AvrHah1, AvrBs3, and AvrBs4 in the proposed recognition domain of AvrBs4 (N-terminal part of the protein up until repeat 3.5) reveals that AvrBs3 and AvrBs4 are 98% identical, but AvrHah1 and AvrBs3 are 89% identical (Fig. 6-3) (66). AvrHah1 and AvrBs3 partially overlap in their DNA binding capabilities, however AvrHah1 is structurally rare in the TAL effector family, as it possesses both 34 and 35 amino acid (aa) repeats in its DBD: repeats 1-6 and 10-12 have 35 aa, while repeats 7-9 and 13 have 34 aa (27). AvrBs4, AvrBs3, Hax3, and Hax4 have 34 aa repeats in the DBD and are all recognized by Bs4. Hax2 is a TAL effector with 35 aa repeats that is not recognized by Bs4 (67). Assuming there is a direct interaction between Bs4 and DBD of TAL effectors, an intriguing hypothesis is that the 35 aa repeat structure allows Hax2 and AvrHah1 to avoid Bs4 recognition.

Although binding of a AvrBs4 deletion variant and Bs4 was not detected in a yeast two hybrid assays, a direct ligand-binding model is consistent with the observation that a truncated (and inactive) version of AvrBs4 containing only the first 3.5 repeats of the DBD maintains recognition by Bs4 (66). Recent work on *Xo1* resistance in rice suggests a Bs4-like mechanism of TAL effector recognition, where TAL effector deletions of all but 3.5 repeats in the DBD are sufficient to trigger resistance (75). Interestingly, replacement of the DBD of tal1C from Xoo with the DBD of AvrHah1 did not activate *Xo1* resistance, whereas a replacement of the RVDs of PthXo1 to reflect those of AvrHah1 resulted in activation of *Xo1*, indicating the importance of DBD structure, but not RVD composition, for *Xo1* resistance (75). Additional experiments using the LRR domain of Bs4 and a more complete minimal recognition domain of TALEs with 35 aa or 34 aa may provide additional information about the possibility of direct ligand-binding.

The dTALEs we constructed are in an AvrBs3 backbone, however we do not observe a HR in tomato when delivered by XgΔAvrHah1. Because a wet apoplast can prevent the development of HR (55), the water soaking activities induced by the dTALE targets may be abrogating the Bs4 resistance response. Although we did not observe a significant increase in apoplastic fluid in the case of dT 514, we also did not see a distinct cell death response. It is possible that the dT 514 targeted pectinesterase induces a low level water soaking effect that has the ability to prevent Bs4-mediated cell death.

7. Investigation of delivery of eukaryotic "effectors" via the Type III Secretion System into plant cells

Background

Although many examples of different types of Type III effectors have been identified, it is unclear what types or sizes of proteins are capable of being secreted via the Type III apparatus. If an understanding of the protein features that are permissive for secretion could be obtained, the Type III secretion system could be utilized as a tool for heterologous protein delivery in plants. The delivery of eukaryotic proteins into animals cells via the bacterial Type III secretion apparatus was recently achieved by fusing the N-terminal fragment of the *Yersinia enterocolitica* secreted protein YopE onto target proteins (76). To our knowledge, a similar approach has not been tried in plants. Delivery of eukaryotic proteins into plant cells via the Type III secretion apparatus would be an alternative tool to *Agrobacterium* transient expression, which can be inefficient in certain host plants such as tomato.

We present a proof of principal experiment where we show that the Type III Secretion System of *Xanthomonas gardneri* can deliver GFP, a eukaryotic protein, into plant cells. We then show the utility of this approach by delivering two tomato genes, the pectate lyase and pectinesterase (both indirect targets of AvrHah1) into tomato and assaying for their effects on water soaking.

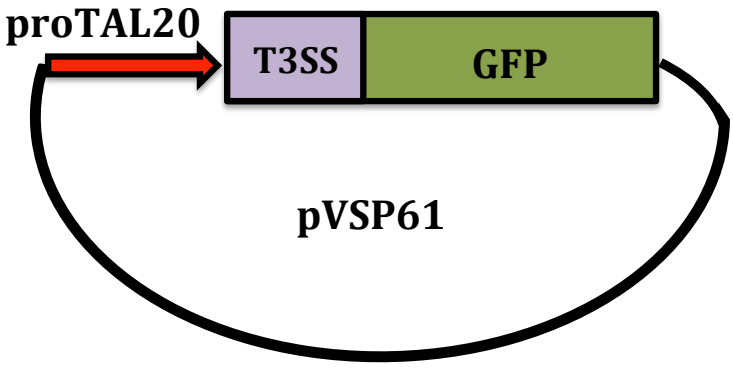


Figure 7-1. Schematic of eukaryotic T3E delivery
The Type III Secretion Signal (T3SS) of AvrHah1 was fused in frame with GFP to create T3E-GFP. The construct is driven by the promoter of TAL20 (proTAL20).

Results

Proof of principal experiment showing Type III delivery of a GFP "effector"

We conducted a proof of principal experiment to determine if the delivery of a eukaryotic protein into plant cells could be accomplished via the Type III Secretion System of Xg. We chose GFP as the eukaryotic protein because it is able to localize to the nucleus of plant cells and can be easily visualized using fluorescence microscopy. To create a eukaryotic Type III effector (T3E) for GFP, we used Gibson assembly to clone together the first 50 amino acids of AvrHah1 (including the ATG) in frame with the coding sequence of GFP (without the ATG) in pVSP61 driven by the TAL20 promoter (Fig. 7-1). The first 50 amino acids of AvrHah1 contain the Type III Secretion Signal (T3SS) of the effector. We transformed this construct into XgΔAvrHah1 and inoculated XgΔAvrHah1 + T3E-GFP into pepper ECW. At 24hpi, we infiltrated pepper leaves with DAPI, a fluorescent stain for DNA that shows the localization of nuclei. We imaged plant leaves using filters for DAPI and GFP and merged the images together to identify if co-localization of nuclei occurred (Fig. 7-2). Indeed, we observed co-localization of nuclei for DAPI and GFP merged images, indicating that the T3E-GFP was successfully translocated via the Xg Type III Secretion System. In contrast, no GFP signal was observed for XgΔhrcV + T3E-GFP, as this mutant does not make a functional Type III Secretion apparatus and thus does not secrete any effectors.

Delivery of AvrHah1 targets as Type III effectors

In Section 5 we described how a dTALE activating Solyc05g014000, the pectate lyase indirect target of AvrHah1, complemented water soaking in XgΔAvrHah1. We hypothesized that successful delivery of the pectate lyase as a Type III effector (T3E) would complement water soaking in XgΔAvrHah1. Similar to TS3-GFP, we cloned the first 50 amino acids of AvrHah1 to the beginning of the cDNA transcript of the pectate lyase. For comparison, we also created a mutant version of the pectate lyase with an Arginine to Alanine mutation in the proposed catalytic domain (R274A) (102). We observed increased water soaking in tomato infected with the effector pectate lyase (PL), XgΔAvrHah1 + T3E-PL, compared to XgΔAvrHah1 alone (Fig. 7-3). We did not observe increased water soaking in response to delivery of the catalytically inactive pectate lyase effector, T3E-PL_{R274A}. These results show evidence that the pectate lyase was delivered into plant cells and capable of inducing water soaking.

We made T3E versions of the AvrHah1-direct target bHLH transcription factors: T3E-$bHLH_{03}$ (Solyc03g097820) and T3E-$bHLH_{06}$ (Solyc06g072520). Because we showed that dTALEs for the two bHLH transcription factors activated expression of the pectate lyase and pectinesterase, we hypothesized that we would see similar gene activation in response to successful delivery of the T3E-bHLHs. However, we did not observe activation of the pectate lyase or pectinesterase (Fig. 7-4). Additionally, we did not observe water soaking (data not shown). These results indicate that the T3E-bHLH transcription factors

were likely not delivered successfully into plant cells or were delivered in a mis-folded configuration that disrupted normal function.

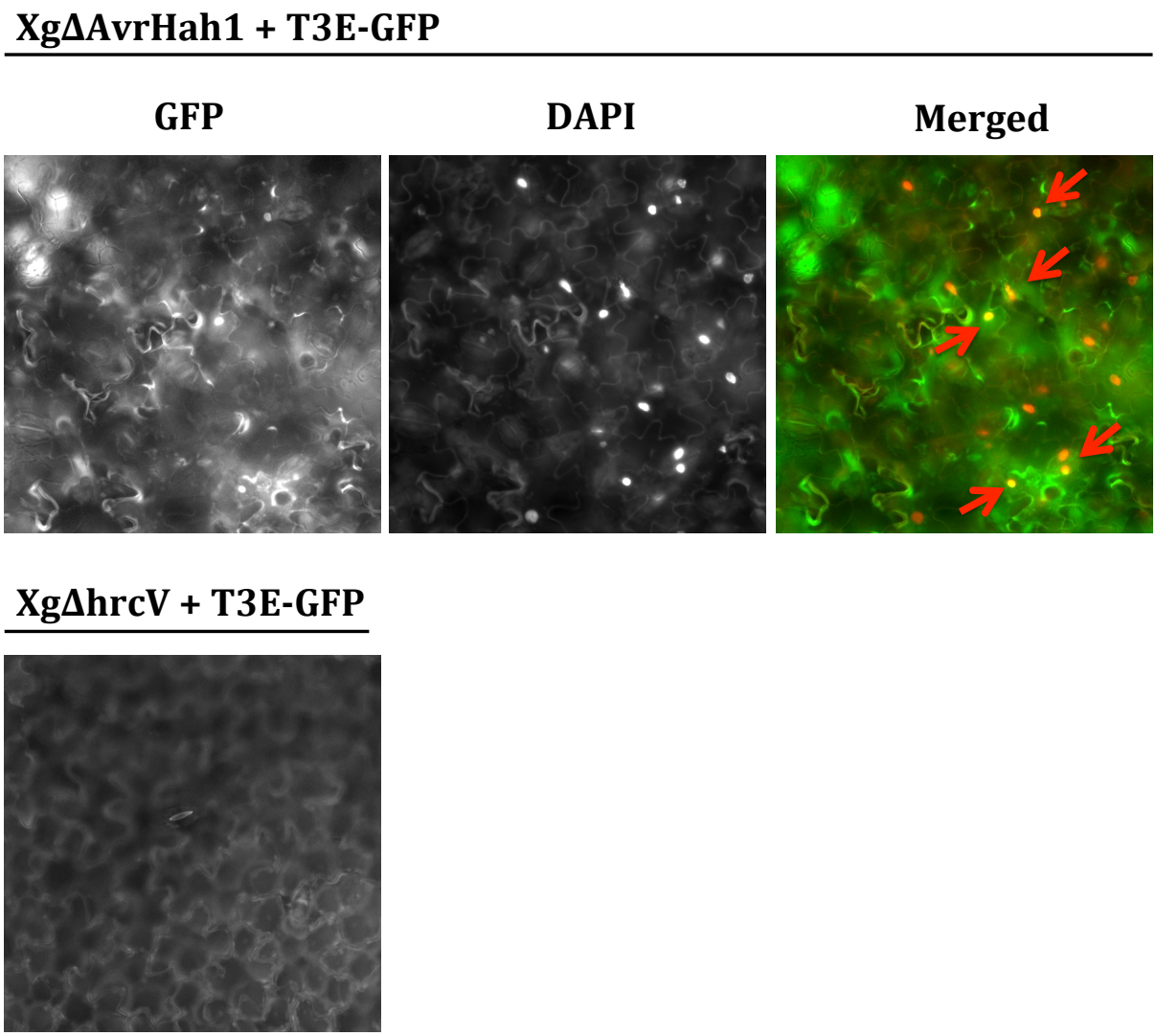


Figure 7-2. Proof of principal experiment demonstrates successful delivery of GFP via the Type III Secretion System
XgΔAvrHah1 + T3E-GFP was infiltrated into pepper ECW at OD_{600} = 0.25. At 24 hpi, leaves were infiltrated with DAPI and imaged for GFP and DAPI. Red arrows indicate merged images of nuclear signal (yellow nuclei) from GFP and DAPI fluorescence. Nuclei with only DAPI signal are colored red in the merged image. No GFP fluorescence is observed in XgΔhrcV + T3E-GFP.

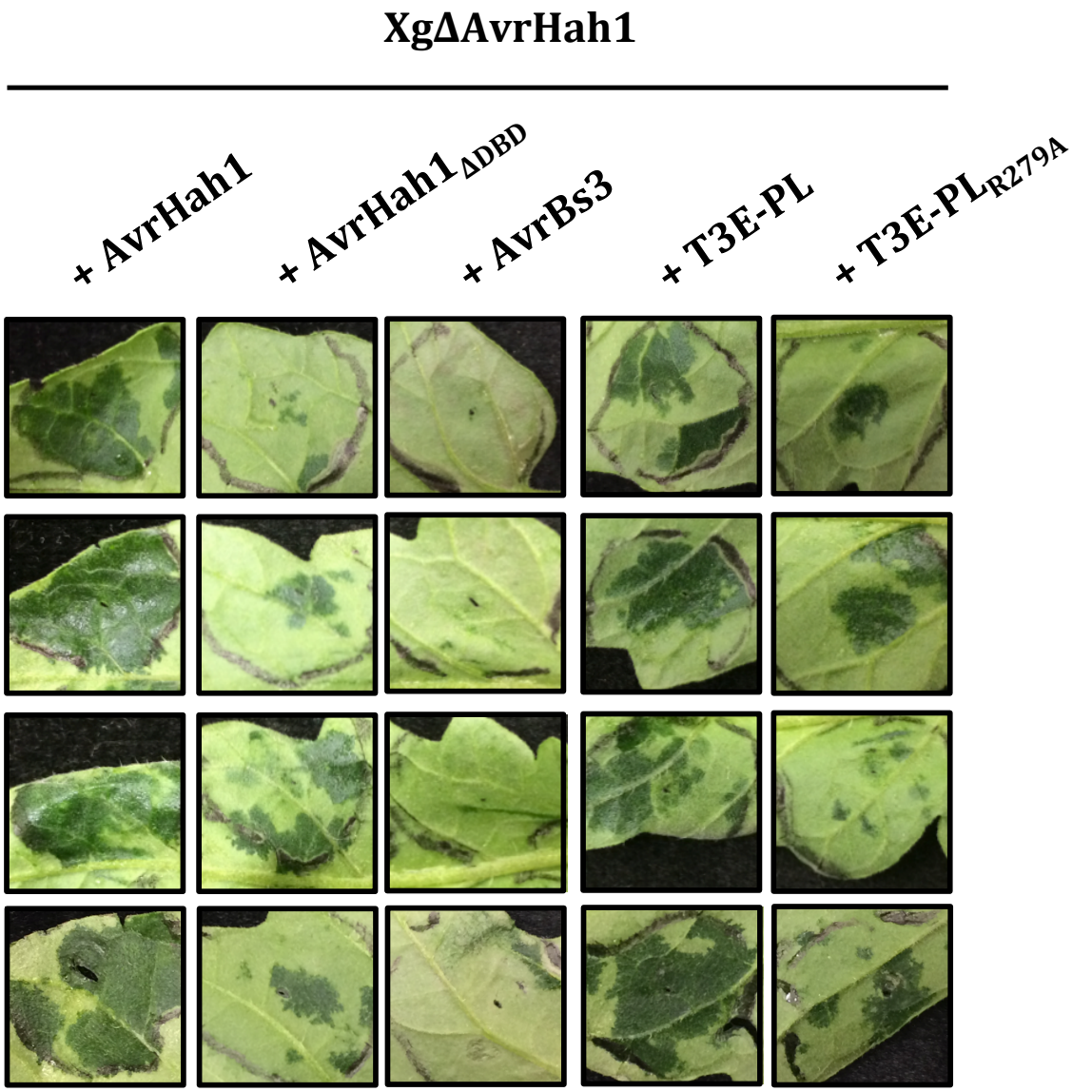


Figure 7-3. Delivery of T3E-Pectate lyase, but not a catalytic inactive mutant, confers partial water soaking to XgΔAvrHah1
T3E-Pectate lyase (PL) and a catalytic mutant T3E-PL_{R274A} were delivered into tomato by XgΔAvrHah1. At 48hpi leaves were submerged in water for 20 minutes to induce water soaking (OD_{600} = 0.1).

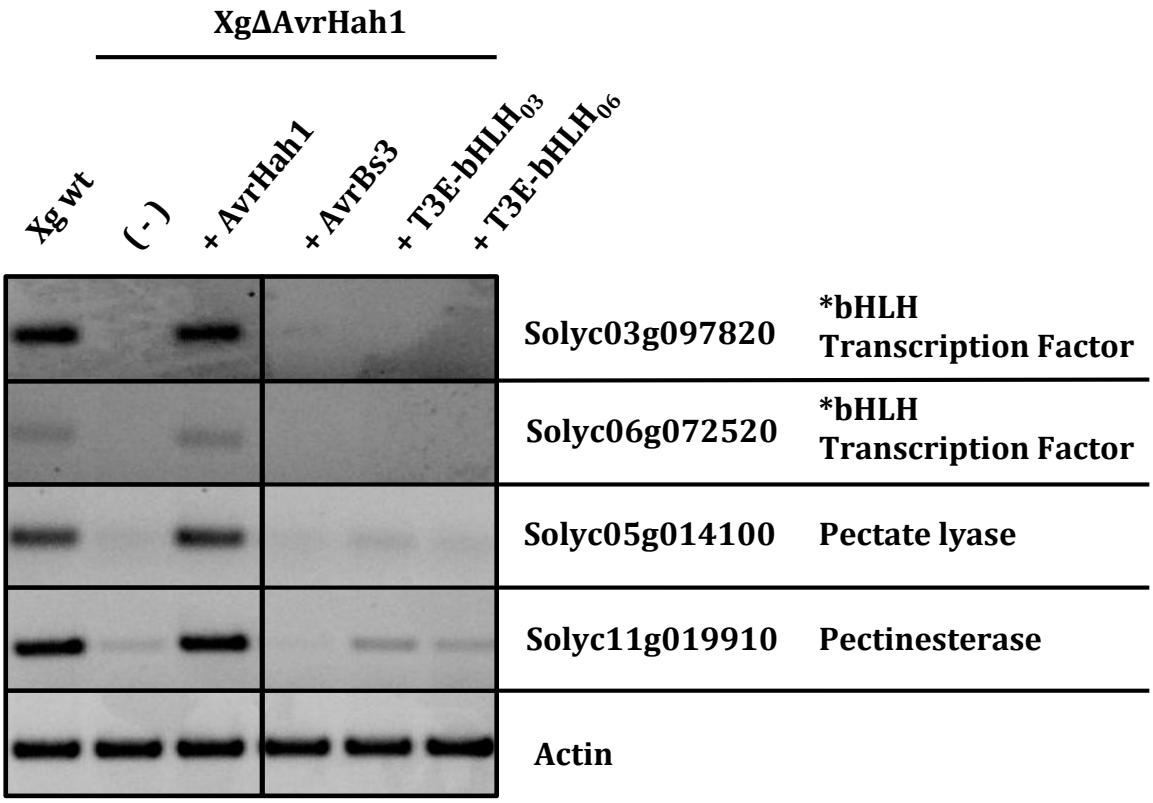


Figure 7-4. Delivery of T3E bHLH transcription factors does not activate expression of target genes

Semi-quantitative RT-PCR of genes on the right performed on tomato Heinz 1706 infiltrated with XgΔAvrHah1 carrying the TAL effectors or T3E-bHLHs (OD_{600} = 0.25, tissue for RNA extraction was collected 24hpi). *Indicates direct target of AvrHah1.

Discussion

We provide evidence of successful delivery of GFP, a eukaryotic protein, into plant cells via the bacterial Type III Secretion System. We accomplished this by fusing the Type III Secretion Signal of AvrHah1 to GFP, driven by the promoter of an effector. We delivered the tomato pectate lyase Solyc05g014000 as a Type III effector via XgΔAvrHah1 and we were able to observe enhanced water soaking compared to delivery of a catalytically inactive version of the pectate lyase. This proof of principal experiment shows the potential utility of heterologous protein delivery via the Type III Secretion apparatus. This method may provide an advantage to *Agrobacterium* transient expression when experiments require biologically relevant levels of protein. Additionally, the speed at which proteins are delivered via the Type III secretion apparatus is likely faster than when synthesized *de novo* in the case of *Agrobacterium* transient expression.

We were not able to recapitulate the expected functional activity of the bHLH transcription factors using this method, indicating that some proteins are likely not amenable to Type III secretion. It may be possible to manipulate protein sequences to become optimized for Type III secretion. As our approach did not contain a positive control for successful translocation, we cannot make any conclusions about at what stage the T3E-bHLHs failed. Development of a positive control for translocation with a screen consisting of variable types and sizes of eukaryotic proteins for their abilities to be secreted would provide additional information on any protein motifs or patterns that could predict the likelihood of successful delivery. Combinations of Type III Secretion signals from various effectors and their promoters could be explored to determine their effects on secretion.

8. Materials and Methods

Xanthomonas strain collection

Xe, Xp, and Xg strains were collected from diseased tomatoes and peppers in the United States. Xp strains were collected between 1998 and 2013 in Florida and Georgia. Xg strains were collected in Ohio and Michigan between 2010 and 2012. Xe strains were collected between 1994 and 2012 in Florida, North Carolina, Georgia, and Kentucky. A database of strains was curated in the Staskawicz lab with unique identification numbers for each strain.

Plant materials

Experiments with pepper were performed using Early CalWonder (ECW), ECW20R (with *Bs2* introgression), and ECW30R (with *Bs3* introgression). Experiments in tomato were performed in Heinz 1706 (Tomato Genetics Resource Center). Experiments in *Nicotiana benthamiana* were performed in the Staskawicz lab's variety "Nb-Seq".

Genome sequencing and effector predictions

Bacterial genome sequencing and effector predictions were completed as previously described (41). Briefly, genomic DNA was isolated with a modified CTAB protocol and prepared for library construction and sequencing on the Illumina platforms. Ten Xg libraries were pooled into a single lane of MiSeq (PE250). Xe and the Xp strains from 2006 were sequenced by multiplexing 48 libraries per lane on an Illumina HiSeq 2000 sequencer (PE100). The Xp strains from 2012 were sequenced by multiplexing 20 libraries per lane on an Illumina MiSeq (PE150). Genomic *de novo* assemblies were constructed using CLC Genomics Workbench using a length fraction of 0.9 and a similarity of 1.0. Potential effectors were identified by an in-house Python script utilizing BLAST against a database of known effectors, using a filter of greater than 45% amino acid similarity over 80% of the length of the target sequence (41).

Phylogenomic inference using core protein-coding genes

All genomes sequenced in this study were annotated using the National Center for Biotechnology Information (NCBI) Prokaryotic Genome Annotation Pipeline (PGAP) Ortholog families were determined using the GET_HOMOLOGUES package, which includes a step of all-against-all BlastP (77) followed by clustering based on OrthoMCL to yield homologous gene clusters (78). This result was filtered using compare_cluster.pl (a script in the GET_HOMOLOGUES package) with option "-t n", where *n* is the number of genomes, keeping only the gene families that have exactly one representative from each genome

considered; the protein-coding genes in these families were considered the 'core genome' of these species.

Accuracy checking of each individual gene alignment (using nucleotide sequences) was performed by Guidance (79) using the Mafft algorithm (80) anchored by codons with default options, followed by the removal of low-accuracy alignment sites. All edited alignments were concatenated by FASconCAT yielding a nucleotide supermatrix (81). The best partitioning scheme and evolutionary model for each partition were calculated by PartitionFinder (82), which tests all available models under the Bayesian Information Criterion (BIC) selection procedure (83) Maximum likelihood (ML) analysis for phylogeny construction was performed using IQTree v.1.1.5 assuming the best partitioning and respective models according to the previous step (84). A total of 1,000 bootstrap pseudoreplicates were performed to assess clade support. Additional taxa included to strengthen the confidence in the phylogenetic relationships are as follows: *Xanthomonas fragariae* (XfrLMG25863, RefSeq PRJNA80793: (85), *Xanthomonas arboricola* pv. *corylina* (XacNCCB100457, RefSeq PRJNA193452: (86), *Xanthomonas campestris* pv. *musacearum* (XcmNCPB4384, RefSeq PRJNA73881: (87), *Xanthomonas axonopodis* pv. *citrumelo* F1 (XalfaF1, RefSeq PRJNA73179: (88), *Xanthomonas oryzae* pv. *oryzae* (XooKACC10331, RefSeq PRJNA12931: (89), *Xanthomonas campestris* pv. *campestris* (XccATCC33913, RefSeq PRJNA57887: (90), *Xanthomonas euvesicatoria* (also *Xanthomonas campestris* pv. *vesicatoria*, Xe85-10, RefSeq PRJNA58321: (91).

Whole genome SNP analysis

Illumina reads were trimmed using Trimmomatic version 0.32 (92) and were then mapped to the reference genome *Xanthomonas axonopodis* pv. *citri* strain 306 (Xac306, NC_003919: (90) using bowtie2 version 2.1.0 (93). The Best Practices guidelines of the Broad Institute for variant calling were followed. MarkDuplicates from Picard Tools version 1.118 was used to mark duplicate reads. RealignerTargetCreator and IndelRealigner from GenomeAnalysisToolkit (GATK) version 3.3-0 were used to verify reads were aligned properly (94). HaplotypeCaller from GATK was used to discover variants. SNPs were concatenated as previously described (41). A ML phylogenetic tree with bootstrap values was created using RAxML version 8.0 (95).

Effector allele analysis

Effectors were compared within each species at the amino acid sequence level for Xp and the nucleotide level for Xe and Xg, and each distinct allele was assigned a number identifier. Neighbor-joining trees were constructed to visualize differences in effector profiles among strains in each species. Simple genetic distances among strains in their effector profiles were calculated for all pairwise comparisons within each species, such that a difference at one effector between two strains equaled a distance of 1.0 and a difference at five effectors equaled a distance of 5.0. Xp calculations included an outgroup profile from Xe85-10. Distance was calculated using GenAlEx 6.501. Distance matrices were exported to MEGA format and trees were constructed in MEGA 6.06 (96).

Confirmation of the TAL effector AvrHah1 in *Xg*

Xg strains were infiltrated into ECW30R at OD_{600} = 0.3 in order to determine if activation of the *Bs3* resistance gene occurs in response to AvrHah1. Negative and positive controls for AvrHah1 are Xg strain 1782 and 04T5, respectively (27). Pictures were taken 48 hours post infiltration (hpi). For Southern blot analysis, 5µg of *Xg* DNA (extracted as described above) was restriction digested for 2 hours with BamHI and run on a 0.7% agarose gel. DNA was transferred overnight to a Hybond-N+ membrane and hybridized overnight with a P^{32}-labeled probe for the first 705bp of AvrHah1. The size of the predicted BamHI-digested AvrHah1 fragment is 2,964bp.

Construction of bacterial mutants

Insertion mutants in Xp strains (ΩavrBsT) were constructed using site-directed homologous recombination of a partial fragment linked to a gene for antibiotic resistance. Intragenetic partial fragments (approximately 500 bp) of each targeted gene were PCR amplified and cloned using the TA cloning method (Invitrogen). The plasmids were introduced into competent cells of Xp recipient strains by electroporation, and transformed cells were selected for kanamycin resistance (kan^R). Single homologous recombination events disrupted the gene of interest (97). Mutations were confirmed by PCR and Sanger sequencing.

All Xg mutants were constructed in Xg153 and all Xe mutants were constructed in Xe85-10. Whole gene knockout strains XeΔXopQ, XgΔhrcV, Xp4BΔAvrBsT, Xp4BΔXopQΔAvrBsT, and XgΔAvrBs2 were constructed using the suicide vector pLVC18 containing the contiguous 1kb upstream and 1kb downstream fragments flanking the targeted gene (58). Double homologous recombination events resulted in markerless deletions. Mutants were confirmed with PCR and Southern blot. Gene deletions of *avrBsT* and/or *xopQ* in Xp and Xe were complemented by conjugation of the stable broad host range plasmid pVSP61 (kan^R) containing the native promoter driving the open reading frame of AvrBsT and/or XopQ, as appropriate.

The XgΔAvrHah1 mutant was created by double homologous recombination using the suicide vector pLVC18 (58), such that the entire 14.5 repeats of the central DNA binding domain (DBD) was deleted in-frame, preserving the N and C terminal thirds of the protein. Deletion of the DBD was confirmed by Southern blot analysis and loss of 30R HR (*Bs3* activation).

Complementation of XgΔAvrHah1

All complementation constructs for XgΔAvrHah1 are driven by 1kb of the TAL20 promoter (proTAL20) (37). proTAL20 and AvrHah1, AvrBs3, dTALEs, and the eukaryotic "type III effectors" were Gibson™ cloned (New England Biosciences) into a $gentamycin^R$ entry vector (98) using *SalI* and *XbaI* sites and Gateway™ cloned (Invitrogen) into the broad host-range vector pVSP61 using LR Clonase (Invitrogen). Triparental matings of complementation plasmids into XgΔAvrHah1 were performed with the *E. coli* helper strain

pRK600, selected for on rifampicin and kanamycin, and confirmed with PCR. dTALEs targeting the promoters of Solyc03g097820, Solyc06g072520, Solyc05g014000, and Solyc11g019910 for activation were constructed as previously described (32).

To create eukaryotic Type III effectors, a DNA sequence for the first 50 amino acids of AvrHah1 (containing the Type III Secretion Signal) was cloned upstream of a cDNA transcript (without the ATG) of the target gene, e.g. GFP, Solyc03g097820, Solyc06g072520, and Solyc05g014000.

Bacterial growth conditions

Xanthomonas strains were grown on nutrient yeast glycerol agar (NYGA) supplemented, as appropriate, with 100 µg/ml rifampicin (all strains) and 25 µg/ml kanamycin (for strains complemented with the pVSP61 vector), and 10 µg /mL tetracycline (for Xg Tet^R used in water soaking inoculum and XpΩavrBsT). Strains were incubated at 28°C for 48 hours. Cells were adjusted to appropriate concentrations with 10mM $MgCl_2$.

Bacterial growth assays

For *in planta* growth assays, leaves were syringe-infiltrated with bacterial suspensions of 10^5 CFU/mL and leaf discs were collected at the times indicated (at time 0 and either 2 dpi or 6 dpi). One 0.5cm^2 leaf discs was ground in 500 µl 10mM $MgCl_2$ for 2 minutes using 2 3mm glass beads and a mechanical beater. For virulence scoring of a HR or water soaking, leaves were syringe infiltrated at 10^8 CFU/mL (OD_{600} = 0.1) and observed 48 hours post-infiltration (hpi) after submerging leaves in water for 20 minutes. For development of discrete lesions, leaves were syringe infiltrated at 10^4 CFU/mL and observed 6 dpi. Surface populations were estimated by plating of a dilution series of water droplets allowed to rest on the leaf surface for 20 minutes.

Water soaking and bacterial introduction assays

Leaves were syringe infiltrated with a bacterial suspension adjusted to OD_{600} = 0.1 (~10^8 CFU/mL). At 48 hours post infection (hpi), water soaking was induced in *N. benthamiana* by creating a small epidermal wound in the infected area and pipetting a 30ul drop of water or 10^5 CFU/mL Xg Tet^R on top of the wound. Total *in planta* bacteria were selected for on NYGA with rifampicin (Rif), and Xg Tet^R internalized during the water soaking were selected by plating on rifampicin and tetracycline (Rif + Tet). For tomato water soaking assays, infected leaves were submerged in water for 20 minutes.

Apoplastic fluid measurements

For quantitative water soaking assays, tomato leaves were syringe infiltrated with a bacterial inoculum adjusted to OD_{600} = 0.1. At 48 hpi, leaves were submerged in water for

20 minutes and blotted with a kimwipe to remove water from the leaf surface. Two 0.5cm^2 leaf discs were collected and placed in a 0.5mL tube with a small hole cut in the bottom. This tube was placed in a pre-weighed 1.5mL tube. The tubes were centrifuged at 8,000 rpm for 5 minutes to collect the apoplastic fluid. Weights of the 1.5mL tubes post-spin were subtracted from the pre-spin weights to obtain a quantitative measurement of apoplastic fluid.

RNA-seq and TAL effector prediction

Xg wt and XgΔAvrHah1 were syringe infiltrated into tomato Heinz 1706 at OD_{600} = 0.25 and tissue was collected and frozen in liquid nitrogen at 24 hpi and 48 hpi. RNA from three biological replicates per time point were prepared using the Spectrum™ Total Plant RNA Kit (Sigma-Aldrich) and sequencing libraries were prepared using the Illumina TruSeq RNA Library Prep Kit v2. Sequencing of 100bp read length with Paired Ends was performed on a single lane of an Illumina HiSeq2000. Data analysis was performed using the CLC Genomics Workbench software to identify differentially expressed genes (at least 2-fold different with $p < 0.05$).

Computational predictions for AvrHah1 EBEs was performed using the TALE-NT 2.0 algorithm (64) with RVDs for AvrHah1 (NN IG NI NI NI HD HD NG NN NI HD HD HD NG) and a cutoff of 4 times the best possible score (3.76) in a Heinz 1706" tomato "promoterome", consisting of the 300bp sequences upstream of all annotated genes.

Semi-quantitative RT-PCR

Xg strains were infiltrated into tomato leaves at OD = 0.25 and tissue was collected at 24 hpi. The Spectrum™ Plant RNA kit (Sigma-Aldrich) (with on-column DNAseI treatment) and the SuperScript™ III First-Strand Synthesis System were used to make cDNA from 1.5 µg of RNA (quantified using the NanoDrop). Five microliters of 1:10 diluted cDNA were used for 24 cycles of amplification using Phusion® HF polymerase (New England Biolabs). Ten microliters of each reaction were used for gel electrophoresis.

Transient promoter::luciferase assays in *N. benthamiana*

A promoter sequence of 1kb upstream of the start codon of Solyc03g097820, Solyc06g072520, Solyc05g014000, and Solyc11g019910 was Gateway™ (Invitrogen) cloned into the binary luciferase reporter construct pGWB35 (99). AvrHah1 and the coding regions of Solyc03g097820 and Solyc06g072520 were cloned into the binary expression vector p1776. All constructs were conjugated into *Agrobacterium* GV3101 via triparental matings. For each combination of promoter and transcriptional activator, leaves of *N. benthamiana* were co-infiltrated with *Agrobacterium* (OD_{600} = 0.4 for each strain). At 24 hpi, the leaves were syringe infiltrated with 1 mM luciferin. Six 0.28 cm^2 leaf punches per condition were taken and placed in separate wells of a black microtiter plate, suspended on 100 µl of water. Luciferase activity was read using a Wallace Envision plate reader.

9. Supporting Information

Supplemental Table S1: *Xanthomonas de novo* sequencing and genome assembly statistics

Species	Strain name	Base number	Read number	Genome size	Contig L50	Contig number	Average contig length	Genome coverage	Sequencing platform	NCBI Accession
X. euvesicatoria	Xe073	619622418	6574270	5308180	176620	101	52556	117	HiSeq200 PE 100	JZRY00000000
	Xe074	499393084	5327546	5271601	174842	106	49732	95	HiSeq200 PE 100	JZLM00000000
	Xe075	711500646	7530748	5443723	136606	126	43204	131	HiSeq200 PE 100	JZLN00000000
	Xe076	634996211	6748464	5245668	187535	134	39147	121	HiSeq200 PE 100	JZLO00000000
	Xe077	613059372	6499960	5094970	210541	196	25995	120	HiSeq200 PE 100	JZLP00000000
	Xe078	691650490	7327568	5431856	198355	280	19399	127	HiSeq200 PE 100	JZLQ00000000
	Xe079	696603659	7395350	5286491	176635	80	66081	132	HiSeq200 PE 100	JZLR00000000
	Xe081	720057860	7655232	5512319	197924	569	9688	131	HiSeq200 PE 100	JZLS00000000
	Xe082	673020745	7170916	5325355	222161	125	42603	126	HiSeq200 PE 100	JZLT00000000
	Xe083	665133302	7071854	5374993	187439	129	41667	124	HiSeq200 PE 100	JZLU00000000
	Xe085	700935225	7405554	5321876	163350	116	45878	132	HiSeq200 PE 100	JZLV00000000
	Xe086	619212173	6570582	5392470	187495	174	30991	115	HiSeq200 PE 100	JZLW00000000
	Xe091	197413676	2094156	5312029	141260	116	45793	37	HiSeq200 PE 100	JZLX00000000
	Xe101	675663365	7178230	5377218	163308	148	36333	126	HiSeq200 PE 100	JZLY00000000
	Xe102	651663155	6937562	5321033	159095	124	42912	122	HiSeq200 PE 100	JZLZ00000000
	Xe103	685807898	7284172	5372384	187297	139	38650	128	HiSeq200 PE 100	JZMA00000000
	Xe104	716851850	7593786	5359869	187438	103	52038	134	HiSeq200 PE 100	JZRP00000000
	Xe105	648157693	6858450	5355745	187438	113	47396	121	HiSeq200 PE 100	JZRQ00000000
	Xe106	510807247	5406042	5355786	163306	110	48689	95	HiSeq200 PE 100	JZRR00000000
	Xe107	483930649	5134052	5319326	163306	102	52150	91	HiSeq200 PE 100	JZRS00000000
	Xe108	684847805	7240728	5362642	159114	146	36730	128	HiSeq200 PE 100	JZRT00000000
	Xe109	653087240	6910756	5421456	137244	202	26839	120	HiSeq200 PE 100	JZRU00000000
	Xe110	626536020	6642384	5137420	177767	101	50866	122	HiSeq200 PE 100	JZRV00000000
	Xe111	673020745	7170916	5156822	210555	101	51058	131	HiSeq200 PE 100	JZRW00000000
	Xe112	688852785	7297316	5314098	187444	85	62519	130	HiSeq200 PE 100	JZRX00000000

(Table S1, continued)

Species	Strain name	Base number	Read number	Genome size	Contig L50	Contig number	Average contig length	Genome coverage	Sequencing platform	NCBI Accession
X. perforans	Xp4B	510927000	8515450	5216274	103057	128	40752	97	HiSeq200 PE 100	JZUX00000000
	Xp2010	320869010	1293178	5234672	191273	65	80533	61	MiSeq PE 150	JZVJ00000000
	TB6	645456976	6823788	5231532	278535	52	100606	123	HiSeq200 PE 100	JZWA00000000
	TB9	717503156	7596372	5244610	249937	73	71844	137	HiSeq200 PE 100	JZWB00000000
	TB15	711869951	7574392	5271807	235986	56	94139	135	HiSeq200 PE 100	JZWC00000000
	Xp3-15	882392717	9772542	5462221	236000	89	61373	162	HiSeq200 PE 100	JZUY00000000
	Xp4-20	1150703122	12697816	5291702	322019	58	91236	217	HiSeq200 PE 100	JZUZ00000000
	Xp5-6	1148928316	12785768	5302685	321641	45	117837	217	HiSeq200 PE 100	JZVA00000000
	Xp7-12	966057756	10704340	5123246	241373	55	93150	189	HiSeq200 PE 100	JZVB00000000
	Xp8-16	892760342	9942430	5262802	249686	56	93979	170	HiSeq200 PE 100	JZVC00000000
	Xp9-5	997840661	11062576	5366991	209683	112	47920	186	HiSeq200 PE 100	JZVD00000000
	Xp10-13	2546861708	28104972	5171704	278539	50	103434	492	HiSeq200 PE 100	JZVE00000000
	Xp11-2	972865318	10778350	5359517	246078	44	121807	182	HiSeq200 PE 100	JZVF00000000
	Xp15-11	1216654128	13411430	5313448	266886	47	113052	229	HiSeq200 PE 100	JZVG00000000
	Xp17-12	2365639510	26369818	5185507	302616	52	99721	456	HiSeq200 PE 100	JZVH00000000
	Xp18-15	2482065148	27718790	5353648	376641	42	127468	464	HiSeq200 PE 100	JZVI00000000
	GEV839	235180837	1083060	5445384	117164	119	45760	43	MiSeq PE 150	JZVK00000000
	GEV872	250030537	1167107	5205269	200220	333	15631	48	MiSeq PE 150	JZVL00000000
	GEV893	233916672	1072694	5144963	141320	107	48084	45	MiSeq PE 150	JZVM00000000
	GEV904	195892956	900776	5143643	137900	131	39264	38	MiSeq PE 150	JZVN00000000
	GEV909	235126195	1086572	5153950	198473	84	61357	46	MiSeq PE 150	JZVO00000000
	GEV915	241236850	1109302	5123115	185929	76	67409	47	MiSeq PE 150	JZVP00000000

(Table S1, continued)

Species	Strain name	Base number	Read number	Genome size	Contig L50	Contig number	Average contig length	Genome coverage	Sequencing platform	NCBI Accession
X. perforans	GEV917	178483107	820012	5146551	157228	133	38696	35	MiSeq PE 150	JZVQ00000000
	GEV936	243347472	1126274	5150252	169692	90	57225	47	MiSeq PE 150	JZVR00000000
	GEV940	232941948	1068517	5154749	210471	84	61366	45	MiSeq PE 150	JZVS00000000
	GEV968	258310480	1175317	5151633	191263	83	62068	50	MiSeq PE 150	JZVT00000000
	GEV993	233928947	1075825	5172487	241297	136	38033	45	MiSeq PE 150	JZVU00000000
	GEV1001	249894951	1161721	5186856	244462	62	83659	48	MiSeq PE 150	JZVV00000000
	GEV1026	221921338	1021697	5182328	217437	216	23992	43	MiSeq PE 150	JZVW00000000
	GEV1044	220210692	1007162	5233846	192104	73	71697	42	MiSeq PE 150	JZVX00000000
	GEV1054	237151700	1081262	5236568	241327	73	71734	45	MiSeq PE 150	JZVY00000000
	GEV1063	295016944	1352259	5172138	163796	88	58774	57	MiSeq PE 150	JZVZ00000000
X. gardneri	Xg153	707314366	7471446	5394175	45144	251	21491	131	HiSeq200 PE 100	JZJR00000000
	Xg156	112890220	846494	5313591	43735	201	26436	21	MiSeq PE 250	JZJS00000000
	Xg157	168119123	1972102	5307601	38544	237	22395	32	MiSeq PE 250	JZJT00000000
	Xg159	167531602	1958340	5304768	40532	238	22289	32	MiSeq PE 250	JZJU00000000
	Xg160	231233863	1846316	5336064	49361	180	29645	43	MiSeq PE 250	JZJV00000000
	Xg164	79312020	1058500	5264384	19172	456	11545	15	MiSeq PE 250	JZJW00000000
	Xg165	86780410	1031646	5314003	29193	301	17654	16	MiSeq PE 250	JZJX00000000
	Xg173	695534158	3612768	5345049	61942	159	33617	130	MiSeq PE 250	JZJY00000000
	Xg174	673874226	3815676	5250516	59240	177	29664	128	MiSeq PE 250	JZJZ00000000
	Xg177	99150031	642826	5135275	39449	236	21760	19	MiSeq PE 250	JZKA00000000

10. References

1. Stall RE, et al. (1994) Genetically diverse groups of strains are included in *Xanthomonas campestris* pv *vesicatoria*. *International Journal of Systematic Bacteriology* 44(1):47–53.

2. Vauterin L, et al. (1995) Reclassification of *Xanthomonas*. *International Journal of Systematic Bacteriology* 45(3):472–489.

3. Jones JB, et al. (2000) Systematic analysis of xanthomonads (*Xanthomonas* spp.) associated with pepper and tomato lesions. *International Journal of Systematic and Evolutionary Microbiology*. 50:1211-1219.

4. Jones JB, et al. (2004) Reclassification of the xanthomonads associated with bacterial spot disease of tomato and pepper. *Systematic and Applied Microbiology* 27(6):755–762.

5. Stall RE, et al. (2009) Durability of resistance in tomato and pepper to xanthomonads causing bacterial spot. *Annu Rev Phytopathol* 47:265–284.

6. Potnis N, et al. (2011) Comparative genomics reveals diversity among xanthomonads infecting tomato and pepper. *BMC Genomics* 12:146.

7. Jones JB, et al. (1998) Diversity among xanthomonads pathogenic on pepper and tomato. *Annual Review of Phytopathology* 36:41–58.

8. Timilsina S, et al. (2015) Multilocus sequence analysis of xanthomonads causing bacterial spot of tomato and pepper plants reveals strains generated by recombination among species and recent global spread of *Xanthomonas gardneri*. *Applied and Environmental Microbiology* 81(4):1520–1529.

9. Jones JB, et al. (1998) Evidence for the preemptive nature of tomato race 3 of *Xanthomonas campestris* pv. *vesicatoria* in Florida. *Phytopathology* 88(1):33–38.

10. Tudor-Nelson SM, et al. (2003) Bacteriocin-like substances from tomato race 3 strains of *Xanthomonas campestris* pv. *vesicatoria*. *Phytopathology* 93(11):1415–1421.

11. Hert AP, et al. (2005) Relative importance of bacteriocin-like genes in antagonism of *Xanthomonas perforans* tomato race 3 to *Xanthomonas euvesicatoria* tomato race 1 strains. *Applied and Environmental Microbiology* 71(7):3581–3588.

12. Horvath DMD, et al. (2012) Transgenic resistance confers effective field level control of bacterial spot disease in tomato. *PLoS ONE* 7(8):e42036–e42036.

13. Ma X, et al. (2011) First report of *Xanthomonas gardneri* causing bacterial spot of tomato in Ohio and Michigan. *Plant Disease* 95(12):1584–1584.

14. Parkinson N, et al. (2009) Phylogenetic structure of *Xanthomonas* determined by comparison of *gyrB* sequences. *International Journal Of Systematic and Evolutionary Microbiology* 59(2):264–274.

15. Almeida NF, et al. (2010) PAMDB, A multilocus sequence typing and analysis database and website for plant-associated microbes. *Phytopathology* 100(3):208–215.

16. Hamza AA, et al. (2012) Multilocus sequence analysis and amplified fragment length polymorphism-based characterization of xanthomonads associated with bacterial spot of tomato and pepper and their relatedness to *Xanthomonas* species. *Systematic and Applied Microbiology* 35(3):183–190.

17. Midha S, et al. (2014) Genomic insights into the evolutionary origin of *Xanthomonas axonopodis* pv. *citri* and its ecological relatives. *Applied and Environmental Microbiology* 80(20):6266–6279.

18. Grant SR, et al. (2006) Subterfuge and manipulation: Type III effector proteins of phytopathogenic bacteria. *Annual Review of Microbiology* 60:425–449.

19. Kay S, et al. (2009) How *Xanthomonas* type III effectors manipulate the host plant. *Current Opinion in Microbiology* 12(1):37–43.

20. Mudgett MB. (2005) New insights to the function of phytopathogenic bacterial type III effectors in plants. *Annual Review of Plant Biology* 56(1):509–531.

21. Jones JDG and Dangl JL (2006) The plant immune system. *Nature* 444(7117):323–329.

22. Hajri A, et al. (2009) A “repertoire for repertoire” hypothesis: repertoires of type three effectors are candidate determinants of host specificity in *Xanthomonas*. *PLoS ONE* 4(8):e6632.

23. Boyd LA, et al. (2013) Plant-pathogen interactions: disease resistance in modern agriculture. *Trends in Genetics* 29(4):233–240.

24. Xiao S (2012) Protecting crops from pathogens: novel approaches to an old problem. *Gene Technology1: e103*.

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

26. Quezado-Duval AM, et al. (2014) Outbreaks of bacterial spot caused by *Xanthomonas gardneri* on processing tomato in central-west Brazil. *Plant Disease* 88:157-161.

27. Schornack S, et al. (2008) Characterization of AvrHah1, a novel AvrBs3-like effector from *Xanthomonas gardneri* with virulence and avirulence activity. *New Phytologist* 179(2):546–556.

28. Boch J and Bonas U (2010) Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annual Review of Phytopathology* 48:419–436.

29. Moscou MJ and Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. *Science* 326(5959):1501–1501.

30. Boch J, et al. (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326(5959):1509–1512.

31. White F (2016) *Xanthomonas* and the TAL Effectors: nature's molecular biologist. *Methods Mol Biol* 1338:1–8.

32. Morbitzer R, et al. (2011) Assembly of custom TALE-type DNA binding domains by modular cloning. *Nucleic Acids Research* 39(13):5790–5799.

33. Cohn M, et al. (2016) Comparison of gene activation by two TAL effectors from *Xanthomonas axonopodis* pv. *manihotis* reveals candidate host susceptibility genes in cassava. *Molecular Plant Pathology* 17(6):875-89.

34. Cernadas RA, et al. (2014) Code-assisted discovery of TAL effector targets in bacterial leaf streak of rice reveals contrast with bacterial blight and a novel susceptibility gene. *PLoS Pathogens* 10(2):e1003972.

35. Li Z, et al. (2014) A potential disease susceptibility gene CsLOB of citrus is targeted by a major virulence effector PthA of *Xanthomonas citri* subsp. *citri*. *Molecular Plant* 7(5):912–915.

36. White FF, et al. (2009) The type III effectors of *Xanthomonas*. *Molecular Plant Pathology* 10(6):749–766.

37. Cohn M, et al. (2014) *Xanthomonas axonopodis* virulence is promoted by a transcription activator-like effector-mediated induction of a SWEET sugar transporter in cassava. *Molecular Plant Microbe Interactions* 27(11):1186–1198.

38. Yang Y, et al. (1994) Host-specific symptoms and increased release of *Xanthomonas citri* and *X. campestris* pv. *malvacearum* from leaves are determined by the 102-bp tandem repeats of *pthA* and *avrb6*, respectively. *Molecular Plant Microbe Interactions* 7(3):345-355.

39. Kay S, et al. (2007) A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science* 318(5850):648–651.

40. Wichmann G and Bergelson J (2004) Effector genes of *Xanthamonas axonopodis* pv. *vesicatoria* promote transmission and enhance other fitness traits in the field. *Genetics* 166(2):693–706.

41. Bart R, et al. (2012) High-throughput genomic sequencing of cassava bacterial blight strains identifies conserved effectors to target for durable resistance. *Proceedings of the National Academy of Sciences* 109(28): E1972-E1979.

42. Minsavage GV, et al. (1990) Gene-for-gene relationships specifying disease resistance in *Xanthomonas campestris* pv *vesicatoria* - pepper interactions. *Molecular Plant Microbe Interactions* 3(1):41–47.

43. Monteil CL, et al. (2013) Nonagricultural reservoirs contribute to emergence and evolution of *Pseudomonas syringae* crop pathogens. *New Phytologist* 199(3):800–811.

44. Astua-Monge G, et al. (2000) Resistance of tomato and pepper to T3 strains of *Xanthomonas campestris* pv. *vesicatoria* is specified by a plant-inducible avirulence gene. *Molecular Plant Microbe Interactions* 13(9):911–921.

45. Swords K, et al. (1996) Spontaneous and induced mutations in a single open reading frame alter both virulence and avirulence in *Xanthomonas campestris* pv *vesicatoria* avrBs2. *Journal of Bacteriology* 178(15):4661-4669.

46. Wichmann G, et al. (2005) Reduced genetic variation occurs among genes of the highly clonal plant pathogen *Xanthomonas axonopodis* pv. *vesicatoria*, including the effector gene avrBs2. *Applied and Environmental Microbiology* 71(5):2418–2432.

47. Kim NH, et al. (2010) *Xanthomonas campestris* pv. *vesicatoria* effector AvrBsT induces cell death in pepper, but suppresses defense responses in tomato. *Molecular Plant Microbe Interactions* 23(8):1069–1082.

48. Wei C-F, et al. (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(1):32–46.

49. Witek K, et al. (2016) Accelerated cloning of a potato late blight-resistance gene using RenSeq and SMRT sequencing. *Nature Biotechnology* 34(6):656-660.

50. Scholthof K-BG (2006) The disease triangle: pathogens, the environment and society. *Nature Reviews Microbiology* 5(2):152–156.

51. Freeman BC and Beattie GA (2009) Bacterial growth restriction during host resistance to *Pseudomonas syringae* is associated with leaf water loss and localized cessation of vascular activity in *Arabidopsis thaliana*. *Molecular Plant Microbe Interactions* 22(7):857–867.

52. Beattie GA (2011) Water relations in the interaction of foliar bacterial pathogens with plants. *Annual Review of Phytopathology* 49(1):533–555.

53. Oh HS 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. *Plant Journal* 44(2):348–359.

54. Wang BHLH TRANSCRIPTION FACTORS, et al. (2009) Analysis of temperature modulation of plant defense against biotrophic microbes. *Molecular Plant Microbe Interactions* 22(5):498–506.

55. Schornack S, et al. (2013) Engineering plant disease resistance based on TAL effectors. *Annual Review of Phytopathology* 51(1):383–406.

56. Romer P, et al. (2007) Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. *Science* 318(5850):645–648.

57. Tai TH, et al. (1999) Expression of the Bs2 pepper gene confers resistance to bacterial spot disease in tomato. *Proceedings of the National Academy of Sciences* 96(24):14153–14158.

58. Lindgren PB, et al. (1986) Gene-cluster of *Pseudomonas syringae* pv. *phaseolicola* controls pathogenicity of bean plants and hypersensitivity on non-host plants. *Journal of Bacteriology* 168(2):512–522.

59. Hartmann N and Büttner D (2013) The inner membrane protein HrcV from *Xanthomonas* spp. is involved in substrate docking during type III secretion. *Molecular Plant Microbe Interactions* 26(10):1176–1189.

60. Schwartz AR, et al. (2015) Phylogenomics of *Xanthomonas* field strains infecting pepper and tomato reveals diversity in effector repertoires and identifies determinants of host specificity. *Frontiers in Microbiology* 6:535.

61. Potnis N, et al. (2015) Plant pathogen induced water soaking promotes *Salmonella enterica* growth on tomato leaves. *Applied and Environmental Microbiology* 81(23):8126-34.

62. Marois E, et al. (2002) The *Xanthomonas* type III effector protein AvrBs3 modulates plant gene expression and induces cell hypertrophy in the susceptible host. *Molecular Plant Microbe Interactions* 15(7):637–646.

63. Kay S, et al. (2009) Detailed analysis of the DNA recognition motifs of the *Xanthomonas* type III effectors AvrBs3 and AvrBs3Δrep16. *The Plant Journal* 59(6):859–871.

64. Doyle EL, et al. (2012) TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. *Nucleic Acids Research* 40(W1):W117–W122.

65. Yang B, et al. (2006) Os8N3 is a host disease-susceptibility gene for bacterial blight of rice. *Proceedings of the National Academy of Sciences* 103(27):10503–10508.

66. Schornack S, et al. (2004) The tomato resistance protein Bs4 is a predicted non-nuclear TIR-NB-LRR protein that mediates defense responses to severely truncated derivatives of AvrBs4 and overexpressed AvrBs3. *Plant Journal* 37(1):46–60.

67. Kay S, et al. (2005) Characterization of AvrBs3-like effectors from a *Brassicaceae* pathogen reveals virulence and avirulence activities and a protein with a novel repeat architecture. *Molecular Plant Microbe Interactions* 18(8):838–848.

68. Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485(7400):635–641.

69. Grau J, et al. (2013) Computational predictions provide insights into the biology of TAL effector target sites. *PLoS Computational Biology* 9(3): e1002962.

70. Sun H, et al. (2015) Genome-wide identification and characterization of the bHLH gene family in tomato. *BMC Genomics* 16:9.

71. Gong P, et al. (2010) Transcriptional profiles of drought-responsive genes in modulating transcription signal transduction, and biochemical pathways in tomato. *Journal of Experimental Botany* 61(13):3563–3575.

72. Lionetti V, et al. (2012) Methyl esterification of pectin plays a role during plant-pathogen interactions and affects plant resistance to diseases. *Journal of Plant Physiology* 169(16):1623–1630.

73. Hematy K, et al. (2009) Host-pathogen warfare at the plant cell wall. *Current Opinion in Plant Biology* 12(4):406–413.

74. Schornack S, et al. (2005) Expression Levels of avrBs3-Like Genes Affect Recognition Specificity in Tomato Bs4- But Not in Pepper Bs3-Mediated Perception. *Molecular Plant Microbe Interac*tions 18(11):1215–1255.

75. Triplett LR, et al. (2016) A resistance locus in the American heirloom rice variety Carolina Gold Select is triggered by TAL effectors with diverse predicted targets and is effective against African strains of *Xanthomonas oryzae* pv. *oryzicola*. *Plant Journal* 87(5):472-83.

76. Ittig SJ, et al. (2015) A bacterial type III secretion-based protein delivery tool for broad applications in cell biology. *Journal of Cell Biology* 211(4):913–931.

77. Altschul SF, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25(17):3389–3402.

78. Li L, et al. (2003) OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. *Genome Research* 13(9):2178–2189.

79. Penn O, et al. (2010) GUIDANCE: a web server for assessing alignment confidence scores. *Nucleic Acids Research* 38(Web Server issue):W23–8.

80. Katoh K, et al. (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* 30(14):3059–3066.

81. Kück P and Meusemann K (2010) FASconCAT: Convenient handling of data matrices. *Molecular Phylogenetics and Evolution* 56(3):1115–1118.

82. Lanfear R, et al. (2012) Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution* 29(6):1695–1701.

83. Lanfear R, et al. (2014) Selecting optimal partitioning schemes for phylogenomic datasets. *BMC Evol Biol* 14:82.

84. Nguyen LT, et al. (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* 32(1):268–274.

85. Vandroemme J, et al. (2013) Draft genome sequence of *Xanthomonas fragariae* reveals reductive evolution and distinct virulence-related gene content. *BMC Genomics* 14:829.

86. Ibarra Caballero J, et al. (2013) Genome Sequence of *Xanthomonas arboricola* pv. *corylina*, isolated from Turkish filbert in Colorado. *Genome Announcements* 1(3):e00246–13.

87. Wasukira A, et al. (2012) Genome-wide sequencing reveals two major sub-lineages in the genetically monomorphic pathogen *Xanthomonas campestris* pathovar *musacearum*. *Genes (Basel)* 3(3):361–377.

88. Darrasse A, et al. (2013) Genome sequence of *Xanthomonas fuscans* subsp. *fuscans* strain 4834-R reveals that flagellar motility is not a general feature of xanthomonads. *BMC Genomics* 14:761.

89. Lee B-M, et al. (2005) The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice. *Nucleic Acids Research* 33(2):577–586.

90. da Silva ACR, et al. (2002) Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature* 417(6887):459–463.

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

92. Bolger AM, et al. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15):2114–2120.

93. Langmead B and Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9(4):357–359.

94. McKenna A, et al. (2010) The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* 20(9):1297–1303.

95. Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9):1312–1313.

96. Tamura K, et al. (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* 30(12):2725–2729.

97. Sugio A, et al. (2005) Characterization of the hrpF pathogenicity peninsula of *Xanthomonas oryzae* pv. *oryzae*. *Molecular Plant Microbe Interactions* 18(6):546–554.

98. Goritschnig S, et al. (2016) Structurally distinct *Arabidopsis thaliana* NLR immune receptors recognize tandem BHLH TRANSCRIPTION FACTORS domains of an oomycete effector. *New Phytologist* 210(3):984–996.

99. Nakagawa T, et al. (2007) Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. *Journal of Bioscience and Bioengineering* 104(1):34–41.

100. Obradovic A, et al. (2008). Integrated management of tomato bacterial spot. *Integrated Management of Diseases Caused by Fungi, Phytoplasma and Bacteria* 2:211-223.

101. Kearney B and Staskawicz BJ (1990). Widespread distribution and fitness contribution of *Xanthomonas campestris* avirulence gene *avrBs2*. Nature 346:385-386.

102. Kita N, et al. (1996). Differential Effect of Site-directed Mutations in *pelC* on Pectate Lyase Activity, Plant Tissue Maceration, and Elicitor Activity. *The Journal of Biologial Chemistry* 271(43): 26529-26535.