Comparative genomic analyses of *Clavibacter michiganensis* subspecies and characterization of their interactions with nonhost plants

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

You Lu

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advised by Dr. Jane Glazebrook and Dr. Carol A. Ishimaru

July 2015

Acknowledgements

I would like to thank my advisors Dr. Jane Glazebrook and Dr. Carol A. Ishimaru, for their guidance and encouragement throughout this thesis project. I would also like to thank the rest of my advisor committee Dr. Fumiaki Katagiri, Dr. Deborah A. Samac and Dr. Neil E. Olszewski, for serving as reviewers and for their helpful suggestions and collaboration.

I am grateful to all the current and previous members in Glazebrook/Katagiri lab and in Ishimaru lab for all the helps and suggestions provided to me and for making a great and supportive teamwork environment.

I would like to thank Dr. Kevin Dorn for suggestions on how to conduct *de novo* genome assembly on a customized desktop computer, Dr. Kevin Silverstein, Dr. Jon Badalamenti and Dr. Ying Zhang for suggestions on PacBio sequencing and on computational analyses, Dr. Karl-Heinz Gartemann for providing *Cmm* strains and Dr. Ramsey Lewis for providing tobacco seeds.

I appreciate the fellowship provided by Monsanto Co. that supported my graduate study for three years, and travel awards provided by the Plant Biological Sciences Graduate Program that allowed me to attend conferences. Most part of this research project was supported by the Hatch Funding provided by the Minnesota Agricultural Experiment Station.

Dedication

This dissertation is dedicated to my wife Xiaoxue Qiu,

who has been selflessly supporting me over the years.

Abstract

Clavibacter is a group of Gram-positive coryneform bacteria, containing several subspecies of *C. michiganensis* which cause diseases on economically important crops worldwide as well as non-pathogenic plant-associated isolates. In contrast to the well-studied Gram-negative plant pathogenic bacteria, little is known regarding the mechanisms of pathogenesis and host-pathogen interactions in Gram-positive bacteria. Recent advances in methods and technologies have spawned several exciting discoveries and provided new directions for the research of *Clavibacter.* Most significantly, with the continuously decreasing cost and increasing capacities of DNA sequencing technologies, complete bacterial genome sequences can be obtained within a relatively short time at a fraction of the cost of Sanger shotgun sequencing of DNA shotgun libraries. Availability of complete genomes allows comparative genomic analyses to efficiently discover the conserved and unique features within each genome.

This thesis reports the use of a combination of both Illumina sequencing and the PacBio Single-Molecule Real-Time sequencing technologies and obtained high quality *de novo* genome assemblies of three *C. michiganensis* subsp. *insidiosus* (*Cmi*). Comparisons between the genomes of the *Cmi* isolates revealed chromosomal rearrangements and heterogeneity of plasmids. By comparing the genome of *Cmi* to *C. michiganensis* subsp. *michiganensis* (*Cmm*), *C. michiganensis* subsp. *sepedonicus* (*Cms*), *C. michiganensis* subsp. *nebraskensis* (*Cmn*) and a non-pathogenic tomato seed-associated *Clavibacter* strain, a gene cluster potentially involved in biosynthesis of the blue pigment indigoidine was discovered in the genome of *Cmi* and a list of candidate genes that may be involved in pathogen virulence was generated.

Cmm and *Cms* secrete serine proteases that are required for triggering hypersensitive responses (HR) on nonhost plants. Using a transient expression assay targeting the recombinant protein to either apoplast or cytosol of tobacco, the serine proteases ChpG from *Cmm* and Chp-7 from *Cms* were found to be the HR elicitors and were demonstrated to function in the apoplast of tobacco plants. The predicted catalytic

serine residue in Chp-7 was found to be required for triggering an HR, as a point mutation exchanging it for a threonine abolished the HR. In addition, nine species of tobacco were compared for their abilities to recognize ChpG or Chp-7. Both were recognized in *Nicotiana tabacum* and *N. sylvestris*, while ChpG was also recognized by *N. clevelandii* and *N. glutinosa*.

C. michiganensis is non-adapted on *Arabidopsis thaliana*. This model organism was used to characterize the nonhost resistance against *C. michiganensis.* Screening of 34 *A. thaliana* accessions identified 11 accessions that showed a chlorotic response to *Cmm* inoculation. The ability to induce a chlorotic response in Kas-1 was conserved among all the five tested subspecies of *C. michiganensis*. The chlorosis elicitor appears to be heat stable and constitutively present in the culture. The genetics of this chlorotic response in *A. thaliana* was studied using both a recombinant inbred population between Kas-1 and Tsu-1 as well as a F2 population from the same parents. One locus controlling the variation was mapped to a 53 kb region on chromosome 4. Variations were also observed in activation of hormone signaling marker genes, and rates of reducing the *in planta* titer of *Cmm*.

Table of Contents

List of Tables

Chapter II

Chapter III

Chapter IV

List of Figures

Chapter II

Chapter III

Chapter IV

Chapter I. Literature Review

Plants growing in natural environments are constantly exposed to various micro-organisms, including some microbial species known to be pathogenic. Yet most plants cannot be infected by most pathogens, due to the multi-layered plant innate immune system. Though the plant innate immune system functions in surveillance of all types of microbes a plant contacts, including viruses, bacteria, fungi and oomycetes, this study is focused on the interactions between plants and bacteria and most of the examples used here are from studies of bacterial pathogens.

The plant innate immune system

Plants perceive the presence of microbes using two different arrays of immune receptors, constituting the two layers of the plant innate immune system. The immune receptors of the first layer are called pattern recognition receptors (PRRs), which recognize microbe-associated molecular patterns (MAMPs), such as bacterial flagellin, bacterial elongation factor Tu (EF-Tu) and peptidoglycans (PGNs) from bacterial cell walls (Zipfel, 2014). Known plant PRRs are either receptor-like kinases (RLKs) or receptor-like proteins (RLPs), both localized at the plasma membrane. Both RLKs and RLPs contain an extracellular ligand binding domain and a transmembrane domain, while RLKs have a cytoplasmic kinase domain, which RLPs lack. Currently, one of the best characterized bacterial MAMP and corresponding PRR pairs is a 22 amino acid epitope from bacterial flagellin, flg22, recognized by Arabidopsis FLS2. Upon binding flg22, FLS2 forms a heterodimer complex with BAK1, and BIK1 is phosphorylated by BAK1 and dissociates from the complex (Lu et al., 2010). BIK1 in turn phosphorylates downstream signaling targets, activating plant defense responses (Li et al., 2014). These defense responses initiated from recognition of MAMPs by PRRs are called pattern triggered immunity (PTI) (Jones and Dangl, 2006).

The defense responses activated in PTI include Ca^{2+} influx into the cytoplasm, production of reactive oxygen species (ROS), activation of mitogen-activated protein

kinase (MAPKs) cascades, plant hormone signaling, activation of defense gene expression, and callose deposition at the cell wall (Boller and Felix, 2009). It has been proposed that PTI is sufficient to inhibit growth of non-adapted pathogens, and provides a means of broad-spectrum protection to plants.

However, well-adapted plant pathogens have evolved to manipulate various components of the plant innate immune system during the course of defense signaling. In the case of phytopathogenic bacteria, phytotoxins and the type III secretion system (T3SS) are involved in suppressing plant PTI. Several pathovars of *Pseudomonas syringae* are known to produce the phytotoxin coronatine, which interferes with stomatal closure following activation of PTI, thereby inhibiting an effective means of preventing surface pathogens from entering the apoplastic space of leaves (Melotto et al., 2006). Coronatine is structurally similar to jasmonic acid isoleucine (JA-Ile), and can thus suppress salicylic acid (SA) accumulation and the defense signaling mediated by SA (Zheng et al., 2012), by exploiting the antagonism between SA and JA (Spoel and Dong, 2008).

P. syringae and other Gram-negative phytopathogens also use the T3SS to deliver type III effectors (T3Es) that hijack plant immune signaling. Targets of T3Es have been identified and belong to various modules of the plant innate immune system, from the initial MAMP perception to the trafficking of antimicrobial molecules. *P. syringae* pv. *tomato* (*Pto*) DC3000 effector AvrPto targets PRRs FLS2 and EFR in Arabidopsis and inhibits their kinase activities (Xiang et al., 2008). Similarly, FLS2 and EFR are also targeted by the *Pto* effector HopAO1, which has tyrosine phosphatase activity that suppresses the activation of both PRRs (Macho et al., 2014). The downstream MAPK signaling cascades are targeted by the *Pto* effector HopAI1, which dephosphorylates Arabidopsis MPK3, MPK6 and MPK4 on threonine residues (Zhang et al., 2007; Zhang et al., 2012). Some T3Es have evolved to interfere with plant hormone signaling. For example, the *P. syringae* pv. *syringae* effector HopZ1a, which possesses acetyltransferase activity, directly binds soybean JAZ1 and multiple Arabidopsis JAZ proteins and facilitates degradation of JAZ proteins (Jiang et al., 2013). The degradation of JAZ proteins leads to activation of JA signaling and indirect suppression of SA signaling, which may be a functional substitution for coronatine. Some T3Es have evolved to disrupt the function of important plant immune regulators. For example, during PTI,

Arabidopsis RIN4 is phosphorylated on the Ser141 residue, which removes its negative regulation of plant defense signaling, and is phosphorylated on Thr166 by two independently evolved *P. syringae* effectors, AvrB and AvrRpm1, which antagonizes the de-repression effect provided by Ser141 phosphorylation (Chung et al., 2014). All the bacterial effectors characterized to date are delivered into host cells, and targets of those effectors are all localized within the cells.

Though usage of T3Es appears to have equipped pathogens with superior ammunition that demolishes the plant immune system, they can be double-edged swords. Plants have evolved a second layer of immune receptors, designated as disease resistance I proteins, which specifically perceive T3Es either directly or indirectly. This recognition of T3E activates effector triggered immunity (ETI), which constitutes the molecular mechanisms underpinning the "gene-for-gene" concept (Flor, 1971), and is usually accompanied with a hypersensitive response (HR) (Jones and Dangl, 2006). Most of the R proteins characterized to date are nucleotide binding-leucine rich repeat (NB-LRR) proteins, divided into two sub-groups depending on the presence of coiled-coil (CC) or toll/interleukin-1 receptor (TIR) domains at the N termini (Eitas and Dangl, 2010). Direct recognition of effectors is known for only a few examples, such as the recognition of AVR-Pita from the rice blast pathogen *Magnaporthe grisea* by the rice R protein Pi-ta through direct binding in the cytoplasm of the host cells (Jia et al., 2000). Examples of indirect recognition of effectors are more common. It is conceptualized as "the guard hypothesis" that R proteins recognize pathogen effectors by monitoring the modifications of host targets of effectors (Dangl and Jones, 2001). Another class of plant R proteins are RLPs, which have an extracellular LRR domain enabling recognition of apoplastic effectors (Stotz et al., 2014). Cf-2 is an RLP-type R protein mediating Avr2-triggered HR in tomato, by monitoring the conformational change of host apoplastic protease Rcr3 due to the binding with Avr2. RLP-type R proteins all seem to require another RLK protein, SOBIR1, for signaling. All the known characterized pathogen effector proteins recognized by RLP-type R proteins are produced by fungal pathogens, while apoplastic localization has only been confirmed for the effectors produced by the tomato pathogen *Cladosporium fulvum* (Stergiopoulos and de Wit, 2009).

While studying model pathosystems, such as the *Arabidopsis-P. syringae* pv *syringae* DC3000 system (Xin and He, 2013), has enabled insightful analyses of molecular mechanisms of various aspects of plant-microbe interactions, the knowledge may not be directly applicable to understand the pathogenicity of phytopathogenic bacteria that are distantly related to these model organisms. For example, the pathogenicity of *P. syringae* and other related Gram-negative pathogens is highly associated with the presence of a functional T3SS and its repertoire of T3Es. Indeed, identification of Avr-R protein pairs and demonstration of virulence functions of T3Es can well explain the virulence mechanisms and host ranges of many pathogens belonging to this phylogenetic group. However, no T3SS or functionally equivalent secretion systems have been identified in Gram-positive bacteria, including the high guanine and cytosine (G+C) content phylum Actinobacteria and the low G+C content phylum Firmicutes, highlighting the existence of one or more alternative virulence strategies utilized by phytopathogens of these groups.

General features of *Clavibacter michiganensis* and associated plant diseases

Clavibacter michiganensis is a small genus in the *Microbacteriaceae* family in phylum *Actinobacteria*, currently consisting of six recognized subspecies, including *C. michiganensis* subsp. *michiganensis* (*Cmm*), *C. michiganensis* subsp. *sepedonicus* (*Cms*), *C. michiganensis* subsp. *insidiosus* (*Cmi*), *C. michiganensis* subsp. *nebraskanensis* (*Cmn*), *C. michiganensis* subsp. *tessellarius* (*Cmt*), and *C. michiganensis* subsp. *phaseoli* (*Cmp*) (Eichenlaub and Gartemann, 2011; Gonzalez and Trapiello, 2014). Recent isolation of pathogenic *C. michiganensis* from pepper (Yim et al., 2011) and two groups of frequently isolated nonpathogenic tomato seed-associated *C. michiganensis* strains (Zaluga et al., 2013) call for additional taxonomic efforts in the *C. michiganensis* species, and one such project has been proposed (Yasuhara-Bell and Alvarez, 2015). All the subspecies of *C. michiganensis* have coryneform morphology, are aerobic, stain Gram-positive, are non-motile and are non-spore forming (Eichenlaub and Gartemann, 2011).

C. michiganensis, along with several closely related species in the *Microbacteriaceae* family, has received increasing attention due to the significant economic loss from crop diseases caused by this group of pathogens. While *C.*

michiganensis as a species causes diseases on crop species across different families, each subspecies has a relatively narrow host range. *Cmm* causes bacterial wilt and canker of tomato, *Cms* causes ring rot of potato, *Cmi* causes bacterial wilting and stunting of alfalfa, *Cmn* causes Goss's wilt on maize, *Cmt* causes leaf freckles on wheat (Eichenlaub and Gartemann, 2011), and *Cmp* causes bacterial bean leaf yellowing (Gonzalez and Trapiello, 2014). Most of the subspecies of *C. michiganensis* can survive in soil if host plant debris is present, which is believed to be one source of the inoculum. *C. michiganensis* enters host plants through wounds or hydathodes. Once in the host, it colonizes the vasculature and causes systemic infection, including seeds (Eichenlaub and Gartemann, 2011). The seed colonization by *C. michiganensis* creates another major source of inoculum for disease transmission (Tsiantos, 1987; Biddle et al., 1990; NÉMeth et al., 1991).

Genomics of Clavibacter michiganensis subspecies

Whole-genome sequences are available for *Cmm* NCPPB382, *Cms* ATCC33113 and *Cmn* NCPPB2581 (Bentley et al., 2008; Gartemann et al., 2008; Eichenlaub and Gartemann, 2011) and there is a draft genome for *Clavibacter* LMG 26808, a non-pathogenic isolate from tomato seeds (Zaluga et al., 2014). The three complete genomes were obtained using the Sanger sequencing method (Sanger et al., 1977), which produces high quality long reads but has limited throughput and is very expensive. The genome of *Clavibacter* LMG 26808 was sequenced using Illumina HiSeq 2000, a high throughput sequencing platform with short turn-around periods but short read lengths, which complicates the downstream data analyses and makes it almost impossible to assemble the sequences into a complete genome.

All three complete genomes of *C. michiganensis* contain a circular chromosome of approximately 3.3 Mb, encoding roughly 3000 genes. A high level of collinearity was described between the chromosomes of *Cmm* and *Cmn*, where no insertion elements were detected (Eichenlaub and Gartemann, 2011; Zaluga et al., 2014). The genome of *Cms* displayed extensive rearrangements compared to *Cmm*, presumably due to the presence of 106 copies of insertion element IS*1121* (Bentley et al., 2008). Two circular plasmids, pCM1 and pCM2, are present in the genome of *Cmm* NCPPB382 (Gartemann et al.,

2008). One circular plasmid, pCS1, and one linear plasmid, pCSL1 (Brown et al., 2002), are present in the genome of *Cms* ATCC33113 (Bentley et al., 2008). The genome of *Clavibacter* LMG 26808 contains one plasmid with a size similar to pCM2, but the content of this plasmid is different from pCM2 (Zaluga et al., 2014). There is no plasmid detected in the genome of *Cmn* NCPPB2581, although plasmid(s) have been reported in other isolates of *Cmn* (Gross et al., 1979). A pathogenicity island (PAI) was described in a low G+C region in the chromosome of *Cmm* NCPPB382, which encodes many genes proven or suspected to contribute to pathogen colonization or virulence (Gartemann et al., 2008). The PAI was not conserved in the genome of *Cms* ATCC33113, although many PAI-encoded genes have homologs in *Cms*, suggesting the organization of the PAI was disrupted during genomic rearrangements while its contents are likely preserved (Bentley et al., 2008). The PAI-encoded genes do not have homologs in the genome of *Clavibacter* LMG 26808, which may explain its non-pathogenic lifestyle (Zaluga et al., 2014).

Pathogenicity factors of *Clavibacter michiganensis*

Prior to the availability of whole-genome sequences and applications of genetic tools in *C. michiganensis* research, numerous early studies of the virulence mechanisms focused on the properties of exopolysaccharides (EPS) and their roles in wilting symptom development (Strobel and Hess, 1968; Ries and Strobel, 1972; Van Alfen and Turner, 1975). However, lack of a clear correlation with resistance and susceptibility of hosts (Van Den Bulk et al., 1989), as well as lack of support by genetic evidence (Paschke and Van Alfen, 1993), rendered the early speculations on EPS as a virulence factor unsubstantiated. Another suspected virulence mechanism of *C. michiganensis* involved an anion channel-forming activity detected in the culture medium of *C. michiganensis* subsp. *nebraskensis* (Schurholz et al., 1991; Michalke et al., 2001), though its relevance to disease symptom development has not been demonstrated.

With the light of molecular biology shed on *C. michiganensis* research, it was observed that a series of cured derivatives of the wild type virulent strain *Cmm* NCPPB382, lacking either one or both of the native plasmids, displayed reduced virulence on its host plant tomato (Meletzus et al., 1993). The CMM102 strain, containing only pCM2, showed a small delay in wilting symptom development and an

approximately 38% alleviation of host biomass-loss due to the disease. In addition, loss of pCM1 correlates with loss of endocellulase activity, which was later confirmed by the cloning from pCM1 of *celA*, an endo-β-1,4-glucanase of family A1 cellulases, comprised of an N-terminal catalytic domain, a central type Iia bacterial cellulose-binding domain and an unusual C-terminal α-expansin-like domain (Jahr et al., 2000). The CMM101 strain, carrying only pCM1, showed greater delay in wilting symptom development and further alleviation of biomass loss, indicating additional essential virulence factors encoded on this native plasmid. The pCM2-encoded *pat-1* gene was later cloned and found to encode a trypsin family serine protease (Dreier et al., 1997). Pat-1 has one Ser residue in the motif Gly-Asp-Ser-Gly-Gly predicted to be catalytically active, which has been demonstrated to be essential for the activity of Pat-1 in restoring the virulence of CMM100 (Burger et al., 2005). Interestingly, the CMM100 strain, lacking both native plasmids, could not induce any wilting symptom, but could still colonize the host and grow to a similar titer as the parental strain, *Cmm* NCPPB382 (Meletzus et al., 1993). This is likely due to the presence of additional virulence factors in the chromosome, helping the pathogen to suppress the plant immune responses, rather than escaping recognition by the host plant. This is supported by the observation that another non-pathogenic *Cmm* strain, NCPPB3123, triggers plant immune responses, leading to low colonization in host plants (Griesbach et al., 2000). Indeed, with the whole genome sequence of *Cmm* NCPPB382 becoming available, a PAI was identified in a low G+C region in the chromosome (Gartemann et al., 2008). A *Cmm* NCPPB382 derivative strain, CMM30-18, lost the entire PAI, resulting in loss of pathogenicity and ability to colonize the host (Gartemann et al., 2008). The PAI contains open reading frames encoding several genes homologous to *pat-1*, designated as *chromosomal homology to pat-1* (*chp*) genes. Among them, *chpC* is required for *Cmm* colonization in tomato and virulence (Stork et al., 2008). A mutation in *chpC* leads to reduced bacterial titer by two orders of magnitude and no induction of wilt symptoms, possibly an indirect effect of the low titer (Stork et al., 2008). One close homolog of *pat-1* in *Cms*, *chp*-7, has been shown to function in wilt symptom development in potato and eggplant (Nissinen et al., 2009). The *Cmm* NCPPB382 genome also encodes six *ppa* family serine proteases in the PAI, and five additional *ppa* genes dispersed in the chromosome and plasmids (Gartemann et al.,

2008), which is a sign of lineage-specific gene family expansion. Mutations in two of the *ppa* family genes, *ppaA* and *ppaC*, have been reported to reduce virulence of *Cmm* (Eichenlaub and Gartemann, 2011), hinting at their roles in pathogen virulence.

In addition, all the subspecies of *C. michiganensis* secrete plant cell wall degrading enzymes in addition to CelA, such as pectate lyase, polygalacturonase and endoglucanases. These cell wall-degrading enzymes may enable the pathogen to utilize extra carbon sources, and collectively contribute to the canker symptom observed in infected tomatoes and potatoes. Damaged cell wall in the xylem vessels may be responsible for impairment of water transport that ultimately leads to the wilt symptom.

Plant responses to *Clavibacter michiganensis* infection

All accessions in the current collection of cultivated tomato (*Solanum lycopersicum*) are susceptible to *Cmm* infection, but quantitative resistance exists in related wild species of tomato, *S. peruvianum* and *S. habrochaites* (Francis et al., 2001). These provided an excellent resource to identify the defense responses that might contribute to disease resistance to the bacterial canker disease of tomato. The first report used a proteomics approach to discover proteins that were differentially expressed between mock-treated and inoculated samples in a susceptible line and two resistant lines each containing a unique quantitative resistance locus (Coaker et al., 2004). Superoxide dismutase and apoplast peroxidase were differentially expressed over time, and the patterns of hydrogen peroxide accumulation were different among genotypes, highlighting possible involvement of oxidative stress in response to *Cmm* (Coaker et al., 2004). This observation was supported by a later study using a cDNA-amplified fragment length polymorphism (AFLP) method to detect differentially expressed transcripts between the susceptible cultivated tomato and resistant wild tomato species (Lara-Ávila et al., 2011). In both studies, the resolution and sensitivity were not ideal, partially due to the lack of complete genome sequences for the plant species studied. Another study focusing on the susceptible tomato species, *S. lycopersicum*, using a more sensitive Affymetrix Tomato Genome Array method, revealed that ethylene biosynthesis was induced by *Cmm* through specific induction of *ACO1* expression, the expression of an ERF/AP2 family transcription factor was highly induced, and ethylene promoted disease

symptom development (Balaji et al., 2008). Interestingly, induction of *ACO1* expression only occurred in *S. lycopersicum* inoculated with the virulent strain *Cmm* NCPPB 382, but not with the non-pathogenic strain CMM100, which may partially explain the lack of disease symptom development by CMM100 (Savidor et al., 2012). In both studies on *S. lycopersicum*, pathogenesis-related (PR) proteins were induced, indicating the host defense responses were activated, raising the question of whether these antimicrobial proteins are effective against *Cmm* and if *Cmm* can actively detoxify some of the PR proteins. Taken together, these studies suggest that *Cmm* triggers host defense responses upon inoculation, virulent strains of *Cmm* can manipulate the host responses through the actions of secreted proteins, and the mechanisms of how the actions of *Cmm* secreted proteins lead to host gene expression changes are yet to be elucidated.

Research objectives

Our knowledge of the pathogenicity of *C. michiganensis* still has gaps, while availabilities of new technologies and resources enable additional approaches to advance our understanding of this group of phytopathogens. This work studied *C. michiganensis* from three aspects described in the following three objectives. The first objective was to obtain complete genome sequences of three *Cmi* strains and to conduct comparative genomic analyses between subspecies of *C. michiganensis*. To do these, the PacBio single molecule real time sequencing (SMRT) technology was used to obtain high quality whole genome sequences, and several bioinformatics analysis methods were applied to identify the features of the genome of each *C. michiganensis* subspecies. The second objective was to elucidate the molecular mechanisms of *Cms* Chp-7 triggered HR in tobacco. I discovered that Chp-7 is recognized in the apoplast of tobacco leaves and HR is dependent on the predicted proteolytic activity of Chp-7. I also identified the HR elicitor from *Cmm* is ChpG. The third objective was to characterize the nonhost resistance against *C. michiganensis* using the model organism *A. thaliana*. Taking advantage of the rich resources of *A. thaliana*, I was able to analyze the natural variations of different *A. thaliana* accessions in response to *C. michiganensis*, and to dissect the genetic architecture controlling one of the variations through mapping.

Chapter II. Complete genome sequences and comparative genomic analyses of three strains of *Clavibacter michiganensis* subsp. *insidiosus*

Summary

Advances in DNA sequencing technologies allow cost-efficient sequencing of a whole prokaryotic genome with high quality. Comparative genomic analyses between genomes of bacteria from various phylogenetic groups can reveal conserved and distinctive features associated with each genome, and provide a basis for future research directions. Previous comparative analyses have generated valuable insights about different groups of *Clavibacter*, but were based on one genome representative of each group. With more genomes added into such analyses, the reliability and resolution of the comparisons can be improved and novel emergent properties can be captured. In this study, we generated complete genome sequences for three strains of *Clavibacter michiganensis* subsp. *insidiosus*. Comparisons between the three genomes showed chromosomal rearrangements, heterogeneity of plasmids and variations in gene content. A list of genes potentially involved in virulence of phytopathogenic subspecies of *C. michiganensis* was identified through comparisons with the genomes of other subspecies of *C. michiganensis* and a non-pathogenic *Clavibacter*.

Introduction

The genus *Clavibacter* currently consists of only one species, *C. michiganensis*, which is further classified into six subspecies recognized as plant pathogens (Eichenlaub and Gartemann, 2011; Gonzalez and Trapiello, 2014). It has been increasingly recognized that plant-associated non-pathogenic isolates closely related to *C. michiganensis* exist widely in natural environments either as epiphytes or endophytes. The molecular biology of Gram-positive phytopathogenic bacteria, such as *C. michiganensis*, has not been as well characterized as their Gram-negative counterparts, which is partially attributable to the scarce resources available for research. Early studies of *C. michiganensis* subsp. *michiganensis* (*Cmm*) and *C. michiganensis* subsp. *sepedonicus* (*Cms*) relied on naturally occurring or induced variants to infer the roles of the underlying genomic regions (Meletzus et al., 1993), followed by extensive mapping to identify the responsible gene(s) (Dreier et al., 1997; Jahr et al., 2000). Forward-genetic screening of transposon mutant populations provided a powerful alternative approach to isolate genes of interest without the prerequisite of an available whole genome sequence (Kirchner et al., 2001). Though there are successful examples of studies applying this strategy (Nissinen et al., 2009), transposon mutagenesis has its drawbacks, such as biases in insertion sites and high reliance on transformation efficiency. Reverse-genetic approaches can overcome most issues associated with transposon mutagenesis by targeting insertions into genes of interest through homologous recombination, thus creating desirable mutants for further functional characterization (Stork et al., 2008). However, a reverse-genetic approach cannot be applied unless the sequences of genes of interest are known. The advantages of knowing genomic sequences prompted whole genome sequencing of *Cmm* NCPPB 382 (Gartemann et al., 2008) and *Cms* ATCC 33113 (Bentley et al., 2008). The availability of the two complete genomes has since then enabled a series of investigations at the molecular level (Chalupowicz et al., 2010; Savidor et al., 2012; Savidor et al., 2014).

Currently, complete genome sequences are available for one strain of each of three subspecies: the tomato pathogen, *Cmm* NCPPB 382 (Gartemann et al., 2008), the potato pathogen, *Cms* ATCC 33113 (Bentley et al., 2008), and the corn pathogen, *C. michiganensis* subsp. *nebraskensis* NCPPB 2581 (Eichenlaub and Gartemann, 2011). One draft genome of the non-pathogenic *Clavibacter* LMG 26808 has also been

published (Zaluga et al., 2014). The chromosomes of *Cmm* NCPPB 382 and *Cmn* NCPPB 2581 show collinearity over most of their lengths (Zaluga et al., 2014), while the chromosome of *Cms* ATCC 33113 has undergone extensive rearrangement relative to the former two, partially due to transpositions of the insertion element IS*1121*, which is present in many copies (Bentley et al., 2008). Limited by only one complete genome for *Cmm*, *Cms* and *Cmn*, the extent of chromosomal rearrangement between strains within the same subspecies is not known. Comparative genomic analyses also revealed presence or absence of genes encoding known virulence factors (Bentley et al., 2008; Eichenlaub and Gartemann, 2011).

The three complete genomes were all sequenced from DNA shotgun clone libraries using the Sanger chain-terminating sequencing method (Sanger et al., 1977). This approach allows assembly of regions with repetitive elements or large duplications, and produces high-quality fully assembled microbial genomes, but is often expensive and slow. The draft genome of *Clavibacter* LMG 26808 was sequenced using the Illumina MiSeq platform, which is capable of sequencing from amplified sheared genomic DNA with millions of short reads (up to 300 bp) produced in parallel in two days. However, to *de novo* assemble a bacterial genome using short reads is very challenging, because many of the repetitive sequences in bacterial genomes are longer than the read length, making it almost impossible to map the exact location of the reads within the repetitive region. One of the recent technical advances for whole genome sequencing is the PacBio Single-Molecule Real-Time (SMRT) sequencing technology, which can directly sequence DNA samples following the SMRTbell template preparation without the need for amplification and produce very long sequencing reads (up to 30 kb) in a high-throughput fashion (Eid et al., 2009). Direct sequencing without amplification also preserves all the modifications on the original DNA templates, allowing direct detection of DNA methylation (Flusberg et al., 2010).

C. michiganensis subsp. *insidiosus* (*Cmi*) causes bacterial wilt of alfalfa (*Medicago sativa* L.), the most widely cultivated forage legume (Metzler et al., 1997). Bacterial wilt of alfalfa occurs throughout most alfalfa growing regions worldwide. For this reason the pathogen is considered a phytosanitary risk for international seed movement. Previous research on *Cmi* has focused primarily on plant resistance and

diagnostic methods (Samac et al., 1998; Marefat et al., 2007). The lack of a complete genome sequence has hampered progress in understanding molecular mechanisms of *Cmi*-host interactions.

In this study, we utilized a combination of Illumina sequencing and PacBio SMRT sequencing technology (Eid et al., 2009) to generate *de novo* assemblies of the complete genome sequences of three *Cmi* strains. *Cmi* R1-1 and R1-3 were isolated in Minnesota, USA, and ATCC 10253 is the type strain representative of this subspecies. When growing on medium containing glucose, the colonies of R1-1 and R1-3 are mucoid, while ATCC 10253 is not. Genomic comparisons among the three *Cmi* strains revealed common structural features as well as variations in chromosomal organization and gene content. Genomic comparisons with other subspecies of *C. michiganensis* identified a set of candidate genes potentially involved in pathogenicity.

Results

PacBio SMRT sequencing and *de novo* genome assembly of *Clavibacter michiganensis* subsp. *insidiosus* strains R1-1, R1-3 and ATCC 10253

The filtered, high quality sequencing data from the PacBio SMRT RSII system yielded 64,575 reads with a mean read length of 10,424 bp for R1-1; 55,750 reads with a mean read length of 8,828 bp for R1-3; and 62,766 reads with a mean read length of 7,524 bp for ATCC 10253 (Figure 1). The polymerase reads were then used to pre-assemble 40,804 subreads with a mean length of 12,002 bp for R1-1; 48,349 subreads with a mean length of 8,970 bp for R1-3; and 47,528 subreads with a mean length of 8,533 bp for ATCC 10253. The mean fold-coverages obtained were 161.92×: 120.11×; and 117.1× for R1-1, R1-3 and ATCC 10253, respectively. After additional assembly and quality improvements using the sequencing data generated from the Illumina platforms, 43 INDEL errors in the PacBio assembly of R1-1; 89 INDEL errors in R1-3; and 114 INDEL errors in ATCC 10253 were corrected. The resulting genome assemblies of R1-1 and R1-3 both contain four contigs, corresponding to one circular chromosome and three circular plasmids, pCI1, pCI2 and pCI3 (Table 1). However, the genome assembly of ATCC 10253 contains one circular chromosome and only one circular plasmid, pCI1 (Table 1).

General features of the chromosomes of *Clavibacter michiganensis* subsp. *insidiosus* strains R1-1, R1-3 and ATCC 10253

All three *Cmi* genomes contain one circular chromosome of approximately 3.2 Mb in size with a very high G+C content (73%), comparable to the published *C. michiganensis* genomes (Table 1). The zero point of the three chromosomes was assigned at the start codon of the *dnaA* gene, to be consistent with other published *C. michiganensis* genomes. The chromosome of *Cmi* R1-1 contains 2,968 protein coding genes (CDS), two rRNA operons (16S, 23S, 5S), and 46 tRNA genes (Table 2; Figure 2). One tRNA with anticodon TGC coding for alanine is not present in the published genomes of *Cmm*, *Cms* or *Cmn*. Its location in a 3 kb region conserved only in the *Cmi* genomes indicates that it may have been acquired by horizontal gene transfer. The chromosomes of R1-3 and ATCC 10253 have the identical sets of rRNAs and tRNAs as R1-1, while the number of predicted CDS varies due to genomic insertions/deletions, variation in the copy number of insertion elements, and presence/absence of phage (Table 1).

The predicted functions of the CDS were categorized based on the Clusters of Orthologous Groups (COG) terms (Tatusov et al., 2000) as described in the Methods and Materials. In total 2,409 CDS are associated with at least one COG number, and 756 CDS remain hypothetical (Table 3; Figure 2). The CDS involved in metabolism and metabolic pathways were analyzed using the KEGG database. Similar to *Cmm*, *Cmi* lacks enzymes to reduce nitrate, nitrite, sulfate or sulfite, and relies on reduced forms of nitrogen and sulfur. Ammonium can be taken up by CmiR1-1_1451, which encodes a ammonium channel protein, and used for glutamine biosynthesis by CmiR1-1_1416 to 1418. Sulfide can be used for synthesizing cysteine from O-acetyl-L-serine by CmiR1-1_0578, which encodes a cysteine synthase.

Given the high occurrence of IS*1122* in the *Cmi* genome that creates many sites with near identical sequences, the sequences of different strains of *Cmi* might reveal evidence of genomic rearrangements due to recombination between these insertions. To determine if the chromosomes of the three *Cmi* strains are organized in the same order, the sequences of the three chromosomes were aligned (Figure 3). There are eight

recombination breakpoints between the chromosomes of R1-1 and ATCC 10253. Among them, four breakpoints coincide with insertion sites of IS*1122* in both genomes, and four breakpoints coincide with insertion sites in one genome. Between R1-1 and R1-3, all six recombination breakpoints detected have IS*1122* insertions in both genomes. Thus, the chromosomal organizations differ among the three *Cmi* strains, and the relative rearrangements are likely due to homologous recombination events between the IS*1122* sites. Although R1-1 and R1-3 were isolated from infected plants treated with an inoculum made from alfalfa roots that had been removed from a disease resistance nursery. The differences in chromosomal organization, indicate there may be a heterogeneous population of *Cmi* in the field in Minnesota.

The collinearity between the genomes of *Cmi* and other subspecies of *C. michiganensis* was visualized by determining the organization of locally collinear blocks (LCBs) (Figure 4) (Darling et al., 2004). The alignment of the chromosomes was broken down into many LCBs, mostly separated by the presence of insertion elements, as well as rearrangements in the genome of *Cms*. The majority of the LCBs are present in all six aligned genomes, indicating the gene contents of the LCBs are preserved. The orders of the LCBs in the *Cmi* genomes are mostly consistent with the orders in the genomes of *Cmm* and *Cmn*, while the genome of *Cms* showed the most re-ordering of the LCBs. Given that 59 out of 102 copies of the IS elements are located at the boundary of LCBs in *Cms* (Bentley et al., 2008), it seems likely that the high copy number of IS elements in *Cms* increased the frequency of recombination events causing chromosome rearrangements.

Characteristics of plasmids in *Clavibacter michiganensis* subsp. *insidiosus*

Plasmids are widely prevalent in *C. michiganensis*. The published genomes of *Cmm* NCPPB 382 and *Cms* ATCC 33113 both contain two plasmids (Bentley et al., 2008; Gartemann et al., 2008), and a previous report of the presence of plasmids in *C. michiganensis* subspecies revealed that most isolates possess two plasmids while, rarely, some isolates contain three (Gross et al., 1979). In the genomes of *Cmi* R1-1 and R1-3, we detected three plasmids, pCI1, pCI2 and pCI3, with G+C contents slightly lower than that of the chromosome (Table 1). pCI1 and pCI2 were so named for their homologies to

pCM1 and pCS1, and pCM2, respectively (Figure 5). On the other hand, pCI3 is not homologous to any known *C. michiganensis* plasmids. To assess whether pCI1, pCI2 and pCI3 are present in other *Cmi* isolates, PCR primers developed specifically for detecting each of the three *Cmi* plasmids were used to test 30 additional *Cmi* isolates (Table 5). pCI1 is present in all 33 *Cmi* strains. On the other hand, pCI2 and pCI3 are not conserved in all the strains. Simultaneous presence of pCI2 and pCI3 was only detected in strains isolated in Minnesota, USA. The strains ISTA-6 and CS86, which contain pCI1 and pCI2, are clonal descendants of ATCC 10253 maintained in different culture collections. Thus, ATCC 10253, the type strain we obtained from ATCC, seems to have lost pCI2.

All three plasmids in *Cmi* contain genes whose functions are predicted to be involved in conjugal transfer of DNA. For example, pCI2_0012, 0013 and 0015 encode proteins homologous to TrwC/TraI, TraA and TraG, respectively. pCI3_0072 appears to to be a TrwC-like relaxase fused with a AAA domain. Several genes on pCI3 are similar to sequences in the pathogenicity island (PAI) in the genome of *Cmm* NCPPB 382. For example, CmiR1-1_pCI3_0001 and CmiR1-1_pCI3_0003 are homologous to *parX* and its neighboring gene encoding an ATPase. In addition, the *Cmi* homologs of *ppaB1/2* (CmiR1-1_pCI3_0056) and pectate lyase *pelA1* (CmiR1-1_pCI3_0057) are organized the same way as in the *Cmm* PAI.

Genetic disruptions associated with insertion elements and prophage in *Clavibacter michiganensis* subsp. *insidiosus*

The insertion element IS*1122* was detected in all three *Cmi* genomes, and its copy number varied from 22 copies in R1-3 to 38 copies in ATCC 10253 (Table 1). The 22 copies in R1-3 are all present in the same genomic contexts in R1-1, while only 16 out of the 22 are conserved in ATCC 10253. Therefore, only a subset of the insertion positions of IS*1122* are conserved among the three *Cmi* genomes, suggesting that this transposon remains active.

While most of the insertion sites of IS*1122* are in intergenic regions, insertions into an operon or into the coding region of a gene do occur. For example, one copy of IS*1122* (CmiR1-1_0252) is located between two genes encoding for a sugar ABC transporter substrate-binding protein and a sugar ABC transporter permease within one

operon. The insertion may not completely block the expression of the downstream genes within the operon, but it may reduce the abundance of the proteins encoded by the downstream genes. The same insertion exists in the genome of R1-3, but not in ATCC 10253. In another case, one copy of IS*1122* (CmiR1-1_0372) inserted in the middle of a gene encoding a phosphoketolase, resulting in disruption of the open reading frame. This was only observed in the genome of *Cmi* R1-1, as the corresponding phosphoketolase genes in R1-3 and ATCC 10253, CmiR1-3_2696 and CmiATCC_0345, respectively, remain intact.

The chromosome of *Cmi* ATCC 10253 has an approximately 35 kb deletion relative to *Cmi* R1-1 and R1-3, replaced by one copy of IS*1122*, which may be the cause of the large deletion. The genes within the 35 kb region are from CmiR1-1_0031 to CmiR1-1_0060 (Table 4). Most of these genes are present in the other *C. michiganensis* genomes, excluding the possibility that this region was specifically acquired by *Cmi* R1-1 and R1-3.

The genomes of *Cmi* contain regions that resemble a prophage or remnant of previous phage insertion. The first region is a 10 kb low G+C region bordered by CmiR1-1_0856 and CmiR1-1_0865, containing three genes likely of phage origin, including a DNA replication initiation protein, a XerC-like integrase and a PinR-like recombinase. The two genes flanking each border of the putative prophage, CmiR1-1_0855 and CmiR1-1_0866, both encode a NAD(P)H-dependent aldo-keto reductase, with approximately 85% nucleotide sequence similarity, which might be a duplication created during the phage insertion. The second phage-related region is between CmiR1-1_1150 to CmiR1-1_1163. CmiR1-1_1153 is a phage tail tape measure protein, and CmiR1-1_1163 encodes a DNA recombinase. This phage insertion might have disrupted one gene originally encoding a Rossmann fold nucleotide-binding protein with high homology to CMN_01351, as the two flanking genes, CmiR1-1_1149 and CmiR1-1_1164, can be aligned to the N and C terminus of CMN_01351, respectively. Interestingly, this phage insertion is present in the genome of *Cmi* ATCC 10253, but not in *Cmi* R1-3. Another possible phage insertion region conserved in all three *Cmi* genomes is bounded by CmiR1-1_1847 and CmiR1-1_1850, and the DNA sequence within has weak similarity to viral tegument protein. This potential phage insertion has possibly

caused the deletion of three genes, a pair of two-component sensor kinase and response regulator and a D-alanyl-D-alanine carboxypeptidase, within a region that is otherwise collinear with the genomes of *Cmn* NCPPB 2581 and *Cmm* NCPPB 382.

Genes in *Clavibacter michiganensis* subsp. *insidiosus* known to be important in virulence of other subspecies

The known virulence factors of *C. michiganensis* are all secreted proteins. Secreted serine proteases belonging to both the *pat-1* and *ppa* families have been shown to be important for colonization and/or symptom development for *Cmm* and *Cms* (Eichenlaub and Gartemann, 2011). The genome of *Cmi* R1-1 contains four homologs of *pat-1* and five homologs of *ppa* family genes (Figure 6). Based on phylogenetic analyses of their amino acid sequences, all of these homologs in *Cmi* are divergent from members of the *pat-1* and *ppa* families known to be important in virulence on tomato or potato, such as Pat-1, ChpC, PpaA, and PpaC in *Cmm* (Dreier et al., 1997; Burger et al., 2005; Eichenlaub and Gartemann, 2011), and Chp-7 in *Cms* (Nissinen et al., 2009). On the other hand, CmiR1-1_pCI1_0036, CmiR1-1_pCI3_0055 and CmiR1-1_pCI3_0056 clustered with ChpE, PpaJ and PpaB1/2, respectively, of *Cmm* (Figure 6). Though the exact roles of the three serine proteases from *Cmm* are not clear, they were detected in the xylem sap of *Cmm*-infected tomato and their expression was highly inducible during infection or by culturing in medium supplemented with tomato sap (Savidor et al., 2012), suggesting they were involved in interactions with hosts.

Certain cell wall degrading enzymes are secreted proteins known or predicted to affect virulence of *Cmm* and *Cms*. The genome of *Cmi* encodes homologs of two cellulases, CelA and CelB, present in *Cmm* and *Cms*. CmiR1-1_pCI1_0025 encodes a homolog of CelA, with 89% amino acid identity, and CmiR1-1_ 2201 encodes a homolog of CelB, with 91% amino acid identity. All three *Cmi* genomes encode a pectinesterase (CmiR1-1_0242), which seems to be a fusion between its homologs CMN_00249 and CMN_00250 in *Cmn* NCPPB 2581. Pectinesterase is not found in the genomes of *Cmm* and *Cms.* The plasmid pCI3 in the genomes of *Cmi* R1-1 and R1-3 also encodes one pectate lyase gene (CmiR1-1_pCI3_0057), homologous to *pelA1* and *pelA2* in *Cmm*. The genome of *Cmi* contains a pair of beta-xylanases, CmiR1-1_1455 and CmiR1-1_2356,

homologous to *xysB* and *xysA*, respectively, in *Cmm*. Other cell wall degrading enzymes present in the genome of *Cmi* are a polygalacturonase (CmiR1-1_2859), a pair of beta-glucosidases (CmiR1-1_pCI3_0033 and CmiR1-1_pCI3_0037) and an endo-glucanase (CmiR1-1_2678).

Comparative genomic analyses of *Clavibacter michiganensis*

To identify genes that may have functions in specific host-pathogen interactions at the subspecies level, the complete genome sequences of *Cmi* were included in a comparative analysis with other published complete and draft genomes of *Clavibacter*. The pseudogenes in the three *Cmi* genomes are not all conserved. To avoid biases in the downstream comparative analysis, we first generated a pan-genome of *Cmi* (hereafter "*Cmi* PAN"), by adding predicted CDS, which are either absent or pseudo in R1-1 but are intact in R1-3 or ATCC 10253, into the list of predicted CDS of the R1-1 genome. This process added 23 CDS to the existing 3,144 CDS from R1-1. The predicted CDS from the genomes of *Cmi* PAN, *Cmm* NCPPB 382, *Cms* ATCC 33113, *Cmn* NCPPB 2581 and *Clavibacter* LMG 26808 were analyzed. In total, 4611 clusters of orthologous genes were identified using the criteria that orthologous genes must share identical sequence covering at least 70% of the length, with a BLAST alignment E-value smaller than 0.0001, and have equal Pfam domain compositions. While the majority of the clusters contain genes that are single copy in each genome that includes them, some clusters contain genes that appear to be paralogues in the genomes. For example, one such cluster contained genes encoding subtilisin-like serine proteases, which break down into three copies from *Cmm* NCPPB 382 and two copies from each of the other analyzed genomes, adding up to 11 genes in one cluster.

The distribution of the 4611 clusters among the five genomes was further dissected and visualized as a Venn-diagram (Figure 7). Two thousand clusters are conserved among all five genomes, while 348 clusters are unique to *Cmi* PAN, with 269 of them consisting of chromosomal genes and 79 consisting of plasmid-encoded genes. The colonies of *Cmi* are often associated with indigo pigments, a unique characteristic within *C. michiganensis*. An 8.5 kb region consisting of seven genes that may be involved in the biosynthesis of indigoidine was identified among the 348 *Cmi*-unique

clusters. This region is conserved among all the three sequenced strains of *Cmi*. These seven genes include one MarR family transcriptional regulator (CmiR1-1_1123), and the remaining six (CmiR1-1_1117 to 1122) are organized as an operon (Figure 8). CmiR1-1_1122 encodes a pseudouridine-5'-phosphate glycosidase and is homologous to *Dickeya dadantii indA* with 64.7% amino acid sequence similarity. CmiR1-1_1121 encodes a phosphoglycolate phosphatase and is similar to *D. dadantii indB* with 57.4% similarity. CmiR1-1_1120 encodes a phosphoribosyl transferase. CmiR1-1_1119 encodes a transporter, which may export the blue pigment. CmiR1-1_1118 encodes a nonribosomal peptide synthetase (NRPS) of a single module type, similar to *indC* in *D. dadantii* and *bpsA* in *Streptomyces lavendulae*, consisting of a N-terminal adenylation domain, an oxidase domain, a center-located acyl carrier domain and a C-terminal thioesterase domain. CmiR1-1_1118 encodes a phosphopantetheinyl transferase with the center region showing 47.6% similarity with *Vogesella indigofera igiA*.

Many of the genes on pCI3 are unique to *Cmi* among *Clavibacter*. CmiR1-1_pCI3_0015 to 0023 encode genes likely involved in inositol metabolism, and are homologous to *iolT*, *iolJ*, *iolC*, *iolA*, *iolB*, *iolD*, *iolE*, *iolG* and *iolH* of *Bacillus subtilis*, respectively. These genes are transcribed in the same direction and therefore may be organized as one operon. No transcription factors were found in the upstream region of these genes, although one GntR family transcriptional regulator (CmiR1-1_pCI3_0025) is located about 580 bp downstream and is oriented in reverse to the putative *iol* operon.

Cmm and *Cms* share 39 clusters of orthologous genes, including *pat-1*, *chpC* of *Cmm* and *chp-7* of *Cms*, which are known virulence factors on solanaceous hosts. The 80 clusters shared only by *Cmm* and the tomato seed-associated non-pathogenic *Clavibacter* might be important for interacting with tomato. The group of 69 genes present only in pathogenic *C. michiganensis*, but not in the non-pathogenic *Clavibacter* might be candidate genes for general pathogenicity functions (Table 6). This group contains four transcription factors, three sugar polymer degrading enzymes, three sugar or peptide transporters, one phytoene synthase, one chalcone synthase, and a set of enzymes for decoration of oligomers of cell wall components.

Discussion

Availability of three complete genome sequences of *Cmi* enabled comparative genomic analyses to identify potential genes involved in the interactions with its legume host, as well as genes involved in general pathogenicity of *C. michiganensis*. It is also valuable for further mining to detect differences among the subspecies of *C. michiganensis* in order to develop molecular diagnostic tools with precise discrimination capabilities. An earlier effort at *de novo* assembly of the genomes of *Cmi* R1-1, R1-3 and ATCC 10253 using paired-end sequencing reads from Illumina GAIIx and Miseq systems resulted in hundreds of contigs containing ambiguous bases (unpublished), due to the presence of repetitive elements and the known sequence specific errors within GC-rich motifs associated with Illumina platforms (Nakamura et al., 2011; Quail et al., 2012). To obtain complete genome sequences for the three strains of *Cmi*, we utilized PacBio SMRT sequencing for its advantages of long read length and unbiased read quality at GC-rich sequences (Quail et al., 2012; Roberts et al., 2013). The combination of Illumina sequencing and PacBio SMRT sequencing enabled *de novo* assembly of three complete genomes of *Cmi.*

Previous comparative genomic analyses of *Clavibacter* were conducted with one whole genome sequence representative of each subspecies or phylogenetic group (Bentley et al., 2008; Gartemann et al., 2008; Zaluga et al., 2014). While these studies provided valuable insights about each group of *Clavibacter*, some features of the one particular genome may not reflect the situation of most strains of the group. For example, chromosomal rearrangement events could occur randomly between homologous sequences, thus different strains of one group of *Clavibacter* may end up with different organizations of their genomes. Plasmids in *Clavibacter* are also known to display heterogeneity (Gross et al., 1979), therefore the presence/absence variation of plasmids could be expected. These issues can be alleviated with availability of multiple complete genomes for one subspecies/ phylogenetic group of *Clavibacter*. In this study, the characterization of *Cmi* based on three complete genomes indeed revealed differences between strains of *Cmi* while providing better confidence for the common features of *Cmi*.

The chromosomes of *Cmi* R1-1, R1-3 and ATCC 10253 have undergone chromosomal rearrangements likely mediated through homologous recombination between the repetitive sequences of insertion elements. Similar observations have been reported at the subspecies level for *C. michiganensis* (Bentley et al., 2008), but not at the isolate level. The variations of IS*1122* copy number and insertion position among the three genomes of *Cmi* strongly suggest this transposon remains active, which could potentially create many sites for recombination, adding flexibilities to this group of organisms that might enable facile adaptation to new environments.

The genomes of *Cmi* R1-1 and R1-3 possess three different circular plasmids, named as pCI1, pCI2 and pCI3, while the genome of *Cmi* ATCC 10253 only contains pCI1. The heterogeneity of plasmids present in different strains of *Cmi* has been documented previously (Gross et al., 1979). In this study, we developed PCR primers based on the genome sequence of *Cmi* R1-1 to specifically detect the presence of pCI1, pCI2 and pCI3 in an additional 30 isolates of *Cmi* from various geographic origins. We discovered that pCI1 is highly conserved among *Cmi* isolates. All of the 33 tested *Cmi* strains showed positive PCR result for pCI1, indicating this plasmid may provide sufficient fitness advantages to *Cmi* to select against losing it, at least in its natural context. Moreover, pCI1 shares high similarities with pCM1 and pCS1, both of which encode *celA* and are required for virulence, while the tomato seed-associated nonpathogenic *Clavibacter* LMG 26808 does not have a plasmid similar to these three (Zaluga et al., 2014). Altogether, this could suggest pCI1, pCM1 and pCS1 might have derived from similar origins and have contributed to the evolution of an ancestral *Clavibacter* toward a pathogenic lifestyle. On the other hand, though pCI2 and pCI3 are not conserved in *Cmi* strains, some genes encoded by the plasmids may improve fitness of *Cmi* or may be involved in pathogenicity. For example, the predicted inositol metabolism operon encoded by pCI3 likely allows *Cmi* to utilize inositol, which is present in soil in large quantities, as an alternative carbon source.

Serine proteases play important roles in the pathogenicity of *C. michiganensis*. The genes homologous to the *pat-1* family and *ppa* family are detected in the genomes of *Cmi.* Some members of these families seem to have non-redundant functions, since single gene mutants of certain homologs led to measurable decreases in virulence or loss of HR

on nonhost plants. The homologues of *pat-1* and *ppa* families in *Cmi* are quite divergent from the ones in *Cmm* and *Cms* known for virulence functions (Figure 6). This suggests they might have been under selection pressures to target divergent plant proteins and thus may play different roles as those known for *Cmm* and *Cms*. These genes are not clustered in the genome of *Cmi*, in contrast to the organization of *chp* genes in the *Cmm* PAI. Several of these homologs, together with the pectate lyase gene, are located on pCI3 in a pattern similar to the PAI in *Cmm*. Given that the *Cmm* PAI and *Cmi* pCI3 were both likely acquired through horizontal gene transfer, it is possible that they originated from a similar yet to be discovered source. It further raises an interesting possibility that these genes may exist in relatively high frequency in the genomes of microorganisms that associate with plants and coordinately function in promoting microbial fitness during interactions with plants.

Pectinesterase catalyses the de-esterification of pectin, which is a major component of the plant cell wall and synthesized in highly methyl-esterified forms. It is thought that de-esterified pectin is more accessible to pectin degrading enzymes, including pectin lyases and polygalacturonases (Willats et al., 2001). The simultaneous presence of a pectinesterase (CmiR1-1_0242), a polygalacturonase (CmiR1-1_2859) and a pectate lyase (CmiR1-1_pCI3_0057) in *Cmi* suggest it may be more efficient in degrading pectin than *Cmm* and *Cms*, which lack pectinesterases. The *Cmi* pectate lyase, which is encoded on a plasmid, shares 94.0% amino acid sequence identity with *Cmm* PelA1, which is encoded on the pathogenicity island. Interestingly, *Cmn* possesses a different type of pectate lyase (CMN_02654), indicating it was acquired from a different origin than the ones in *Cmi* and *Cmm*.

The two previously described transcriptional regulators, Vatr1 and Vatr2 (Savidor et al., 2014), involved in pathogenicity of *Cmm* are highly conserved among all five analyzed genomes, including the genome of the nonpathogenic isolate LMG 26808. Therefore, while mutants with defects in either gene result in decreased levels of virulence gene induction without affecting bacterial growth in culture (Savidor et al., 2014), it is unlikely that Vatr1 and Vatr2 function only in pathogenicity. It is possible that other genes functioning downstream of Vatr1 or Vatr2 play more direct roles in regulating expression of virulence genes. For example, CMM_0089, which encodes a

CatR LacI-family transcriptional regulator, and CMM_0678, which encodes a PknC family serine/threonine-protein kinase, were both down-regulated by Vatr1 when *Cmm* was grown in an infection-mimicking medium (Savidor et al., 2014), and are only found in the genome of *Cmm*.

The blue pigmentation of *Cmi* colonies is a distinctive feature among *C. michiganensis* and is caused by production of a water insoluble extracellular indigoidine (Starr, 1958). Gene clusters involved in biosynthesis of indigoidine have also been reported in *Dickeya dadantii* (previously *Erwinia chrysanthemi*) (Reverchon et al., 2002), *Streptomyces lavendulae* (Takahashi et al., 2007), *S. chromofuscus* (Yu et al., 2013) and *Vogesella indigofera* (unpublished, GenBank AF088856). Indigoidine is synthesized by the single NRPS gene in each of these identified gene loci. These NRPSs, including CmiR1-1_1118, all share the same set of modules and have an oxidation domain integrated into the N-terminal adenylation domain (Takahashi et al., 2007). However, the organizations of the gene clusters are variable among different bacterial species. The gene cluster in *Cmi* is most similar to the one in *S. chromofuscus* (Yu et al., 2013), both containing one gene encoding a transmembrane transporter and one gene encoding a phosphoribosyl transferase. In addition, *Cmi* has one phosphoglycolate phosphatase gene in this operon. The production of indigoidine could confer tolerance to oxidative stresses and has been shown to promote virulence of *D. dadantii* (Reverchon et al., 2002). It is possible that indigoidine production in *Cmi* also provides similar advantages, and may be a contributing factor in the virulence of *Cmi* in its hosts.

A list of genes conserved in phytopathogenic subspecies of *C. michiganensis* but not in the non-pathogenic tomato seed-associated *Clavibacter* was generated in this study through comparative analyses of the predicted CDS in each genome (Table 6). These genes may play roles in general virulence of phytopathogenic *C. michiganensis*. Other orthologous gene clusters shared by different subspecies of *C. michiganensis* infecting similar hosts may also be useful. For example, the 39 gene clusters that are shared only between *Cmm* and *Cms*, which both infect solanaceous plants, may be important for virulence on these plants through interaction with solanaceous-specific targets. These genes can be considered as candidates for future studies on functional characterization.

Materials and Methods

Bacterial strains and culture conditions

Cmi R1-1 and R1-3 were isolated from stems of *Medicago truncatula* accession R108 infected with inoculum prepared from symptomatic alfalfa roots collected from disease nurseries at the University of Minnesota Research and Outreach Center in Rosemount, Minnesota. *Cmi* ATCC 10253 is the type strain of *Cmi* and was obtained from the American Type Culture Collection. All *Cmi* strains were cultured on YGM medium (De Boer and Copeman, 1980) at 25 °C.

Genomic DNA extraction and whole genome sequencing

Cmi cultures were grown in 50 ml YGM broth to OD_{600}=1.0 and then centrifuged to collect bacterial cells. Preparation of genomic DNA of *Cmi* R1-1, R1-3 and ATCC 10253 for whole genome sequencing was performed with a modified protocol based on the CTAB method developed for Gram-negative bacterial genomic DNA extraction (Wilson, 2001). A step involving incubation with lysozyme (Sigma-Aldrich, St. Louis, MO, USA) at a final concentration of 2.5 μg/ml at 37 °C for one hour was added prior to the lysis step in the original protocol. The resulting DNA pellets were dissolved in 0.5×TE buffer (5 mM Tris-HCl, 0.5 mM EDTA, pH 8.0) and adjusted to concentrations specified by the downstream sequencing facilities.

Cmi R1-1 was sequenced as paired-ends with a 75 bp read length using Illumina GAIIx. R1-3 and ATCC 10253 were sequenced as paired-ends with a 150 bp read length using Illumina MiSeq. All three genomes were also sequenced using the PacBio SMRT sequencing technology, which was performed by the Molecular Biology Core at Mayo Clinic (Rochester, MN, USA). Each sample was prepared as a 20-kb insert library for PacBio P6-C4 chemistry followed by BluePippin size selection at 15 kb cut-off. Each library was then sequenced using one SMRT cell in the PacBio RSII sequencing system. The SMRT sequencing data were quality filtered with a read quality score cut-off of 0.80 in the PacBio SMRT portal (version 2.2.0).

De novo genome assembly

The initial assembly of the PacBio sequencing reads was conducted using the hierarchical genome-assembly process 3 (HGAP3) algorithm (Chin et al., 2013) in the PacBio SMRT portal (version 2.2.0). The minimum subread length cut-off value was set at 7500 bp, 4000 bp and 4000 bp for the genome assembly of R1-1, R1-3 and ATCC 10253, respectively. The resulting assemblies were corrected with the Quiver consensus algorithm (Chin et al., 2013) implemented in the Resequencing Pipeline in the PacBio SMRT portal using the same set of PacBio sequencing data with the default subread length cut-off setting, in order to obtain a high-accuracy genome assembly. Further improvement of the quality of the genome sequences was performed with Pilon (version 1.10) (Walker et al., 2014) using the sequencing data generated by the Illumina GAIIx platform for R1-1, and Illumina MiSeq platform for R1-3 and ATCC 10253. The aligned read files required by Pilon as input were generated using Bowtie2 (version 2.2.4) (Langmead and Salzberg, 2012) and SAMtools (version 1.2) (Li et al., 2009). All the contigs showed greater than 5 kb overlapping sequences at both ends, and were manually closed to circular contigs. These were the final versions of the *de novo* assemblies used in the downstream analyses.

Genome annotation

The genome of *Cmi* R1-1 was annotated with the NCBI Prokaryotic Genome Annotation Pipeline (http://www.ncbi.nlm.nih.gov/genome/annotation_prok/) and the RAST Server (Aziz et al., 2008). Most of the annotation from the two annotation pipelines was consistent, and the conflicts were manually curated based on sequence homologies with the published genomes of *Cmm* NCPPB 382 and *Cms* ATCC 33113, resulting in a consensus version of the annotation of *Cmi* R1-1. The annotation of R1-3 and ATCC 10253 genomes was performed with the RAST Server and manually curated based on the consensus version of the R1-1 genome annotation. Signal peptides for protein secretion were predicted with the SignalP 4.1 server (Petersen et al., 2011).

Functions of the *Cmi* genes were predicted by searching the NCBI Conserved Domain Database (CDD) (http://www.ncbi.nlm.nih.gov/cdd/) with the default parameter setting. The *Cmi* genes associated with a Clusters of Orthologous Groups (COG) hit from the CDD search results were grouped based on the COG functional categories (Tatusov et

al., 2000). Metabolic pathways were analyzed using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (http://www.genome.jp/kegg/). The predicted CDS from *Cmi* PAN was annotated using the BlastKOALA tool (http://www.kegg.jp/blastkoala/) and the resulting KEGG annotation was used for pathway mapping and module mapping with the KEGG Mapper tools (http://www.genome.jp/kegg/mapper.html).

PCR test for *Cmi* plasmids

Presence of *Cmi* plasmids pCI1, pCI2 and pCI3 were tested by colony PCR using primers specifically designed to amplify unique regions on each plasmid. Bacterial cells were boiled in water for 10 minutes and used as PCR templates. pCI1 was detected with primers pCI1_F (5' CTACTCGTTCACGGACTGTGC 3') and pCI1_R (5' GATACCGGCTTGCGACTTCAA 3'), with an expected PCR product size of 663 bp. pCI2 was detected with primers pCI2_F1 (5' GGACCGTCACGAGGATCACTA 3') and pCI2_R1 (5' ACAGCGGTCATCTACGAGGAA 3'), or with pCI2_F2 (5' AGAGGGTCGGCTAGATGTGTC 3') and pCI2_R2 (5' CCAGAACTCGTACTCGGGTCT 3'), with expected PCR product sizes of 506 bp and 611 bp, respectively. Positive amplification with either pair of pCI2 PCR primers was considered positive for the presence of pCI2. pCI3 was detected with primers pCI3_F (5' GATTCTGTTGCCGATCGGTGT 3') and pCI3_R (5' CGTCCACTTGCGCAAAATAGC 3'), with an expected PCR product size of 699 bp. PCR products were analyzed with DNA gel electrophoresis.

Comparative genome analysis

The sequence homology of plasmids was visualized using dot-plots drawn with Dotter (version 4.34.3) (Sonnhammer and Durbin, 1995) with sliding window size set to 200. pCI3 and pCSL1 do not show significant sequence homology to any known plasmids in *C. michiganensis* genomes, thus their dot-plots were not shown. The collinearity between the chromosomes of *Cmi* R1-1, *Cmi* R1-3, *Cmi* ATCC 10253, *Cmm* NCPPB 382, *Cms* ATCC 33113 and *Cmn* NCPPB 2581 was analyzed with Mauve (version 2.4.0) (Darling et al., 2004). The alignment between chromosomal sequences

was performed with the built-in function "progressive Mauve" using the default settings for all parameters except the seed weight, which was set to 20.

To analyze the presence and absence of orthologous genes among the subspecies of *C. michiganensis*, a pan-genome for *Cmi*, designated as *Cmi* PAN, containing all the non-redundant CDS from the genomes of *Cmi* R1-1. R1-3 and ATCC 10253 was created. The CDS from *Cmi* PAN, *Cmm* NCPPB 382, *Cms* ATCC 33113, *Cmn* NCPPB 2581 and *Clavibacter* LMG 26808 were first clustered for orthologous genes using the software package GET_HOMOLOGUES (version 1.3) (Contreras-Moreira and Vinuesa, 2013) with the following parameters set differently from the default: t=0, C=70, E=1e-04, D=1, which stand for reporting all clusters, 70% minimum coverage in BLAST pairwise alignment, E-value cut-off at 1e-04 and equal Pfam domain composition required for consideration as orthologous genes, respectively. Both the COGtriangle algorithm (Kristensen et al., 2010) and the orthoMCL algorithm (Li et al., 2003), which are implemented in GET_HOMOLOGUES, were used to compute the clustering of orthologous genes. The consensus clustering of orthologous genes was determined based on the intersection of initial clustering results from the two algorithms using the script compare_clusters.pl included in GET_HOMOLOGUES, resulting in a matrix listing the genes in each cluster from the input genomes. The information from this matrix was then extracted using the script parse_pangenome_matrix.pl included in GET_HOMOLOGUES, to compute the number of clusters present or absent in each input genome. A Venn diagram was drawn using the R package 'VennDiagram' (version 1.6.9) to visualize the result.

Tables and Legends

Table 1. Summary of general characteristics of the genomes of three strains of *Clavibacter michiganensis* subsp. *insidiosus*

Strain	R1-1	R1-3	ATCC10253
Chromosome	3,207,520 bp, 73.0% G+C	3,193,464 bp, 73.0% G+C	3,186,916 bp, 72.9% G+C
pCI1	47,690 bp, 66.5% G+C	47,291 bp, 67.8% G+C	48,869 bp, 67.7% G+C
pCI2	49,401 bp, 67.6% G+C	48,316 bp, 67.6% G+C	N.D.
pCI3	103,451 bp, 66.2% G+C	102,376 bp, 66.1% G+C	N.D.
CDS	3,143	3,114	3,011
CDS on chromosome	2,968	2,945	2,961
rRNA	6	6	6
tRNA	46	46	46
IS*1122*	27	22	38
Secreted proteins	262	240	227

N.D., not detected. **CDS,** protein coding sequence.

Table 2. tRNAs with particular anticodons in *Clavibacter michiganensis* subsp. *insidiosus*

Amino Acid	Count	Anticodon					
Ala	4	AGC: 0	GGC: 1	CGC: 1	TGC: 2		
Gly	3	ACC: 0	GCC: 1	CCC: 1	TCC: 1		
Pro	3	AGG: 0	GGG: 1	CGG: 1	TGG: 1		
Thr	3	AGT: 0	GGT: 1	CGT: 1	TGT: 1		
Val	3	AAC: 0	GAC: 1	CAC: 1	TAC: 1		
Ser	4	AGA: 0	GGA: 1	CGA: 1	TGA: 1	ACT: 0	GCT: 1
Arg	4	ACG: 1	GCG: 0	CCG: 1	TCG: 0	CCT: 1	TCT: 1
Leu	5	AAG: 0	GAG: 1	CAG: 1	TAG: 1	CAA: 1	TAA: 1
Phe	1	AAA: 0	GAA: 1				
Asn	1	ATT: 0	GTT: 1				
Lys	2			CTT: 1	TTT: 1		
Asp	1	ATC: 0	GTC: 1				
Glu	2			CTC: 1	TTC: 1		
His	1	ATG: 0	GTG: 1				
Gln	2			CTG: 1	TTG: 1		
Ile	1	AAT: 0	GAT: 1		TAT: 0		
Met	3			CAT: 3			
Tyr	1	ATA: 0	GTA: 1				
Cys	1	ACA: 0	GCA: 1				
Trp	1			CCA: 1			

Table 3. COG categories of predicted genes in *Clavibacter michiganensis* subsp. *insidiosus* strain R1-1

COG categories and descriptions	Number of genes[a]
INFORMATION STORAGE AND PROCESSING	
[J] Translation, ribosomal structure and biogenesis	164
[A] RNA processing and modification	0
[K] Transcription	241
[L] Replication, recombination and repair	169
[B] Chromatin structure and dynamics	0
CELLULAR PROCESSES AND SIGNALING	
[D] Cell cycle control, cell division, chromosome partitioning	38
[Y] Nuclear structure	0
[V] Defense mechanisms	32
[T] Signal transduction mechanisms	106
[M] Cell wall/membrane/envelope biogenesis	151
[N] Cell motility	0
[Z] Cytoskeleton	0
[W] Extracellular structures	0
[U] Intracellular trafficking, secretion, and vesicular transport	33
[O] Posttranslational modification, protein turnover, chaperones	79
METABOLISM	
[C] Energy production and conversion	129
[G] Carbohydrate transport and metabolism	272
[E] Amino acid transport and metabolism	239
[F] Nucleotide transport and metabolism	64
[H] Coenzyme transport and metabolism	108
[I] Lipid transport and metabolism	90
[P] Inorganic ion transport and metabolism	142
[Q] Secondary metabolites biosynthesis, transport and catabolism	75
POORLY CHARACTERIZED	

[R] General function prediction only	366
[S] Function unknown	189
NO COG HIT[b]	756

COG annotation was performed using the NCBI Web CD-Search Tool.

[a], Some genes are associated with more than one COG category and are counted in both.

[b], No hit at the default E-value cut-off (0.01).

Table 4. Chromosomal genes present in *C. michiganensis* subsp. *insidiosus* strain R1-1 and absent in *C. michiganensis* subsp. *insidiosus* ATCC 10253

Locus	Identified domain(s)	Predicted gene function
CmiR1-1_0031	pfam12697	Alpha/beta hydrolase family protein
CmiR1-1_0032	pfam12681	Glyoxalase-like domain-containing protein
CmiR1-1_0033	pfam01261	Xylose isomerase-like protein
CmiR1-1_0034	pfam01408; pfam02894	Oxidoreductase
CmiR1-1_0035	pfam12697	Alpha/beta hydrolase family protein
CmiR1-1_0036	pfam13646	HEAT repeats containing protein
CmiR1-1_0037	pfam13411	MerR HTH family regulatory protein
CmiR1-1_0038	pfam01738	Dienelactone hydrolase family protein
CmiR1-1_0039	pfam00440	TetR family regulatory protein
CmiR1-1_0040	pfam02682	ATP-dependent urea carboxylase
CmiR1-1_0041	pfam03007	Condensation domain containing protein
CmiR1-1_0042	N. D.	hypothetical protein
CmiR1-1_0043	N. D.	hypothetical protein
CmiR1-1_0044	N. D.	hypothetical protein
CmiR1-1_0045	pfam02796	Helix-turn-helix domain of resolvase
CmiR1-1_0046	pfam00550; pfam00501; pfam07993	Long-chain acyl-CoA synthetases with a putative dehydrogenase domain
CmiR1-1_0047	N. D.	Putative ketoreductase, subgroup 1
CmiR1-1_0048	pfam13561	Enoyl-(Acyl carrier protein) reductase
CmiR1-1_0049	N. D.	hypothetical protein
CmiR1-1_0050	pfam13796	Putative sensor protein
CmiR1-1_0051	pfam05956	APC basic domain-containing protein
CmiR1-1_0052	pfam01470	Pyroglutamyl peptidase
CmiR1-1_0053	pfam01494	2-polyprenyl-6-methoxyphenol hydroxylase
CmiR1-1_0054	pfam01470	Pyroglutamyl peptidase
CmiR1-1_0055	N. D.	hypothetical protein
CmiR1-1_0056	N. D.	Putative major facilitator superfamily protein
CmiR1-1_0057	N. D.	Putative transcriptional regulator

CmiR1-1_0058	pfam00202	Adenosylmethionine-8-amino-7-oxononanoate aminotransferase
CmiR1-1_0059	pfam00465	Iron-containing alcohol dehydrogenase
CmiR1-1_0060	pfam13649	Methyltransferase

N. D., not detected.

Table 5. Plasmid distribution among strains of *Clavibacter michiganensis* subsp. *insidiosus*

Strain Index	Strain Name	Isolation Location	Year Isolated	pCI 1	pCI 2	pCI 3
ISTA-1	LNPV 0.51	Europe	NA	+	-	-
ISTA-3	NCPPB 1110	USA	1943	+	-	-
ISTA-6	CFBP 2404[a]	USA	1955	+	+	-
048A	048A	California, USA	2012	+	-	+
114	114	California, USA	2012	+	-	+
253	253	California, USA	2012	+	-	-
CS80	DAR26781	New South Wales, Australia	1973	+	-	-
CS82	DAR34844	New South Wales, Australia	1980	+	-	+
CS83	DAR35660a	New South Wales, Australia	1981	+	+	-
CS84	DAR26653/ PDDCC4190	New Zealand	1975	+	+	-
CS85	DAR26656/ PDDCC3570	New Zealand	1972	+	+	-
CS86	DAR26658/ PDDCC2621[b]	USA	1955	+	+	-
CS87	DAR26661/ PDDCC2613	UK	1964	+	+	-
C1-5	C1-5	Minnesota, USA	1994	+	-	+
C1-6	C1-6	Minnesota, USA	1994	+	+	+
C1-7	C1-7	Minnesota, USA	1994	+	+	+
C1-9	C1-9	Minnesota, USA	1994	+	+	+
ND2-1	ND2-1	Minnesota, USA	1995	+	+	+
N3-3	N3-3	Minnesota, USA	1995	+	+	-
N1313	N1313	Minnesota, USA	1995	+	-	+
FL1-20	FL1-20	Minnesota, USA	1995	+	-	+
ND1-6	ND1-6	Minnesota, USA	1995	+	-	+
ND3163	ND3163	Minnesota, USA	1995	+	-	+
ND3142B	ND3142B	Minnesota, USA	1995	+	-	+
ND1-3	ND1-3	Minnesota, USA	1995	+	+	+
FL1-3	FL1-3	Minnesota, USA	1995	+	+	+
N214	N214	Minnesota, USA	1995	+	+	+
R1-1	R1-1	Minnesota, USA	2009	+	+	+
R1-2	R1-2	Minnesota, USA	2009	+	+	+
R1-3	R1-3	Minnesota, USA	2009	+	+	+
R1-4	R1-4	Minnesota, USA	2009	+	+	+
R1-8	R1-8	Minnesota, USA	2009	+	+	+

ATCC10253	ATCC 10253[c]	Kansas, USA	1960	+	-	-

+, detected with PCR; **-,** not detected.

[a] Type strain of *Cmi*, clone of ATCC 10253.

[b] Type strain of *Cmi*, clone of ATCC 10253.

[c] Type strain of *Cmi*, obtained from the American Type Culture Collection.

Table 6. Orthologous genes conserved only in pathogenic subspecies of *Clavibacter michiganensis*

Locus				Predicted Function	Identified Domain(s)
Cmi	Cmm	Cms	Cmn		
CmiR1-1_0074; CmiR1-1_2090	CMM_0145; CMM_2330	CMS_2510; CMS_2800	CMN_00059; CMN_02288	putative secreted lipoprotein	N. D.
CmiR1-1_0075	CMM_0146	CMS_2801	CMN_00060	putative peptidase	N. D.
CmiR1-1_0077	CMM_0148	CMS_2803	CMN_00062	putative sortase-sorted surface-anchored protein	N. D.
CmiR1-1_0092	CMM_0163	CMS_0333	CMN_00078	putative metallo-dependent hydrolase	N. D.
CmiR1-1_0093	CMM_0164	CMS_0334	CMN_00079	hydrolase	N. D.
CmiR1-1_0154	CMM_0199	CMS_0366	CMN_00121	conserved hypothetical protein	N. D.
CmiR1-1_0155	CMM_0200	CMS_0367	CMN_00122	conserved hypothetical protein	N. D.
CmiR1-1_0203	CMM_0247	CMS_0822	CMN_00198	putative oxidoreductase/ dehydrogenase	pfam01408; pfam02894
CmiR1-1_0272	CMM_0316	CMS_0885	CMN_00288	ATPase with chaperone activity	pfam00004; pfam07724; pfam10431
CmiR1-1_0273	CMM_0317	CMS_0891	CMN_00289	alpha-galactosidase	pfam02065
CmiR1-1_0276	CMM_0322	CMS_0894	CMN_00292	putative dipeptidase	pfam01244
CmiR1-1_0277	CMM_0323	CMS_0895	CMN_00293	acetyltransferase	pfam00583
CmiR1-1_0293	CMM_0339	CMS_0900	CMN_00310	conserved membrane protein	pfam07332
CmiR1-1_0294	CMM_0340	CMS_0901	CMN_00311	putative membrane protein	N. D.
CmiR1-1_0295	CMM_0341	CMS_0902	CMN_00312	putative membrane protein	N. D.
CmiR1-1_0296	CMM_0342	CMS_0903	CMN_00313	conserved hypothetical protein	N. D.
CmiR1-1_0298	CMM_0343	CMS_0905	CMN_00315	putative alkaline shock protein	pfam03780
CmiR1-1_0445	CMM_0512	CMS_0425	CMN_00467	putative membrane protein	pfam10990

CmiATCC_0575[a]	CMM_0676	CMS_2145	CMN_00639	dsDNA exonuclease subunit	pfam13476
CmiR1-1_0665	CMM_0740	CMS_2231	CMN_00697	acetyltransferase	pfam02474; pfam13673
CmiR1-1_0677	CMM_0751	CMS_1025	CMN_00706	acyl-CoA dehydrogenase/ methyltransferase	pfam02585; pfam05401
CmiR1-1_0692	CMM_0763	CMS_1040	CMN_00718	GntR family transcriptional regulator	pfam00155; pfam00392
CmiR1-1_2396	CMM_0798	CMS_0057	CMN_00754	conserved hypothetical protein	pfam07944
CmiR1-1_2385	CMM_0809	CMS_0066	CMN_00765	oxidoreductase;	pfam07992
CmiR1-1_2355	CMM_0833	CMS_0088	CMN_00793	nucleotide sugar epimerase	pfam01370
CmiR1-1_2292	CMM_0891	CMS_0149	CMN_00853	aminoglycoside phosphotransferase	pfam04655
CmiR1-1_2222	CMM_0955	CMS_2884	CMN_00917	putative peptide ABC transporter, permease component	N. D.
CmiR1-1_0782	CMM_1037	CMS_0654	CMN_01001	heavy metal-transporting P-type ATPase	pfam00122; pfam00403; pfam00702
CmiR1-1_0797	CMM_1052	CMS_0671	CMN_01016	putative amidophosphoribosyl-transferase	pfam00156
CmiR1-1_0946	CMM_1192	CMS_1899	CMN_01162	ATP-dependant RNA helicase	pfam00270; pfam00271
CmiR1-1_1001	CMM_1243	CMS_0995	CMN_01213	ABC-type xylose transport system, periplasmic component	pfam13407
CmiR1-1_1183	CMM_1398	CMS_0786	CMN_01370	putative alpha-amylase	pfam00128
CmiR1-1_1208	CMM_1422	CMS_1517	CMN_01397	acetyltransferase	pfam00583
CmiR1-1_1255	CMM_1469	CMS_1503	CMN_01446	SAM-dependant methyltransferase	pfam13847
CmiR1-1_1319	CMM_1534	CMS_1773	CMN_01510	chalcone synthase	pfam02797
CmiR1-1_1371	CMM_1589	CMS_1560	CMN_01567	peptide ABC transporter, permease component	pfam00528
CmiR1-1_1406	CMM_1624	CMS_1607	CMN_01604	TetR family transcriptional regulator	pfam00440; pfam02909
CmiR1-1_1495	CMM_1698	CMS_1665	CMN_01677	NAD-dependent epimerase	pfam01370
CmiR1-1_1580	CMM_1784	CMS_2031	CMN_01764	conserved hypothetical protein	N. D.
CmiR1-1_1743	CMM_1945	CMS_1282	CMN_01926	acyl-CoA dehydrogenase	pfam08028
CmiR1-1_1746	CMM_1949	CMS_1280	CMN_01928	MFS permease	pfam00083
CmiR1-1_1773	CMM_1990	CMS_1242	CMN_01952	tyrosyl-tRNA synthetase	pfam00579; pfam01479

CmiR1-1_1774	CMM_1991	CMS_1241	CMN_01953	conserved hypothetical protein	N. D.
CmiR1-1_1775	CMM_1992	CMS_1240	CMN_01954	conserved membrane protein	N. D.
CmiR1-1_1776	CMM_1993	CMS_1239	CMN_01955	putative 3-methyladenine DNA glycosylase	pfam02245
CmiR1-1_1812	CMM_2030	CMS_1203	CMN_01992	Cell division protein	pfam01580; pfam09397; pfam13491
CmiR1-1_1854	CMM_2072	CMS_1161	CMN_02034	DNA glycosylase	pfam00730; pfam06029
CmiR1-1_1895	CMM_2110	CMS_1111	CMN_02075	FAD-dependent oxidase	pfam01565; pfam04030
CmiR1-1_1899	CMM_2114	CMS_1107	CMN_02080	protease II (Oligopeptidase B)	pfam00326; pfam02897
CmiR1-1_1902	CMM_2116	CMS_1102	CMN_02082	thiamine-phosphate pyrophosphorylase/ phosphomethylpyrimidine kinase	pfam02581; pfam03070; pfam08543
CmiR1-1_1935	CMM_2149	CMS_1843	CMN_02116	translation initiation factor IF-2	pfam00009; pfam03144; pfam04760; pfam11987
CmiR1-1_1954	CMM_2168	CMS_2199	CMN_02135	MoxR-like ATPase	pfam07726
CmiR1-1_2032	CMM_2255	CMS_2444	CMN_02230	anti-sigma factor antagonist	pfam01740
CmiR1-1_2039	CMM_2262	CMS_2455	CMN_02237	conserved hypothetical protein	pfam04977
CmiR1-1_2145	CMM_2386	CMS_2323	CMN_02340	secreted tyrosine phosphatase	pfam13348; pfam13350
CmiR1-1_2437	CMM_2460	CMS_2774	CMN_02412	conserved hypothetical protein	pfam10544; pfam10708; pfam13250
CmiR1-1_2462	CMM_2484	CMS_2750	CMN_02436	putative isomerase	pfam13462
CmiR1-1_2514	CMM_2530	CMS_0594	CMN_02487	conserved hypothetical protein	N. D.
CmiR1-1_2561	CMM_2578	CMS_0322	CMN_02534	phosphohexomutase	pfam02880
CmiR1-1_2577	CMM_2593	CMS_0307	CMN_02549	PTS system, phosphocarrier protein, HPr component	pfam00381
CmiR1-1_2578	CMM_2594	CMS_0306	CMN_02550	PTS system, cellobiose-specific component IIB	pfam02302
CmiR1-1_2622	CMM_2637	CMS_0416	CMN_02593	zinc finger domain-containing protein	pfam07336; pfam11706
CmiR1-1_2654	CMM_2669	CMS_2639	CMN_02621	two-component system response regulator	pfam00072; pfam00196
CmiR1-1_2663	CMM_2679	CMS_2650	CMN_02638	conserved hypothetical	N. D.

				protein	
CmiR1-1_2664	CMM_2680	CMS_2651	CMN_02639	MarR family transcriptional regulator	pfam13463
CmiR1-1_2784	CMM_2797	CMS_2611	CMN_02757	alpha-glucosidase	pfam00128; pfam11941
CmiR1-1_2827	CMM_2839	CMS_2957	CMN_02802	NADH(P)-binding domain-containing protein	pfam13460
CmiR1-1_2843	CMM_2854	CMS_2937	CMN_02818	putative ATPase	pfam13614
CmiR1-1_2875	CMM_2887	CMS_2607	CMN_02853	putative bifunctional phytoene synthase/ phytoene dehydrogenase	pfam00494; pfam01593; pfam13450

N. D., not detected.

[a]**,** This gene in R1-1 and R1-3 contains a nonsense mutation from a frameshift, but is intact in ATCC 10253.

Figures and Legends

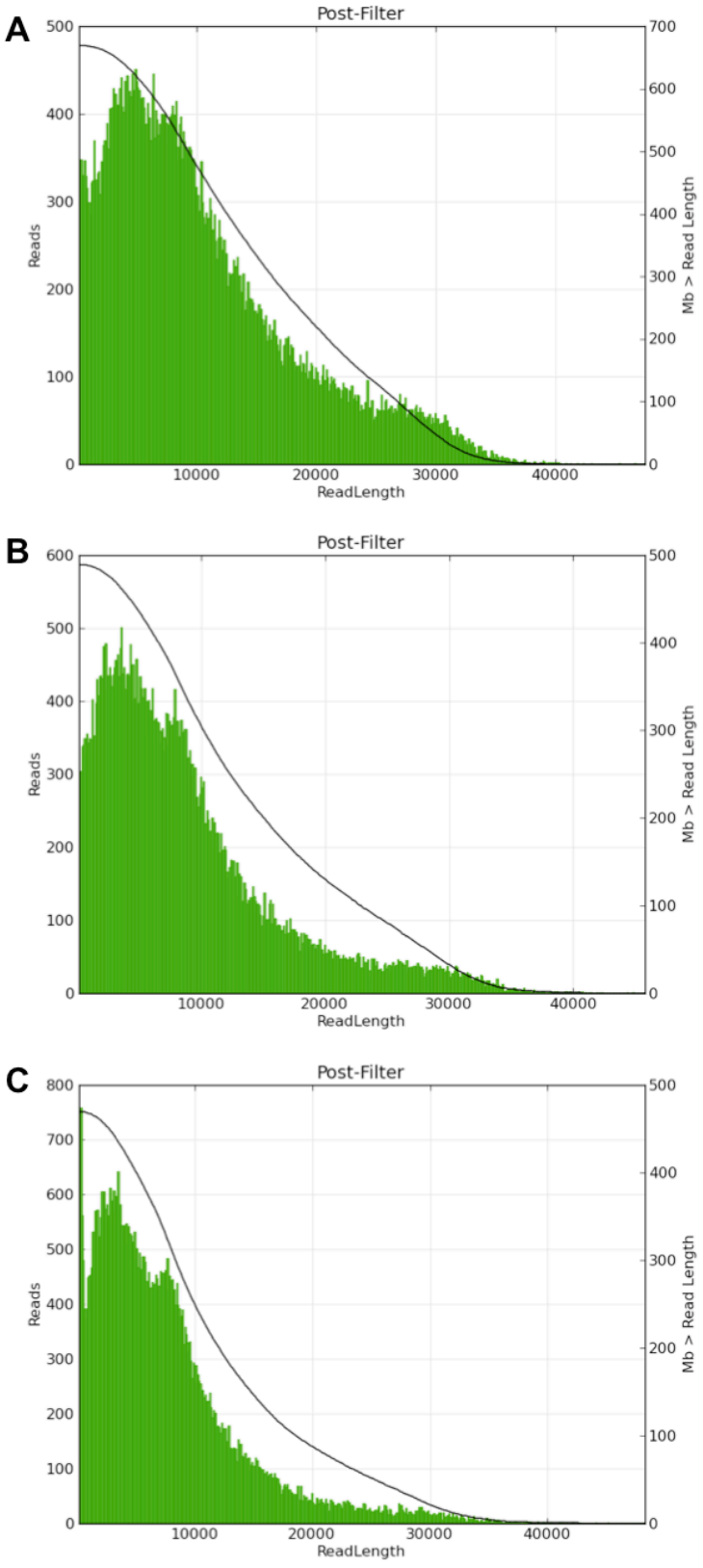


Figure 1. Post-filter polymerase read length and cumulative read length. The polymerase reads with a quality score equal or greater than 0.80 were collected. The number of reads (y-axis on the left) at each read length (x-axis) is plotted and the cumulative read length (y-axis on the right) is calculated starting from the longest polymerase read. The plots are from the sequencing data of **A,** R1-1; **B,** R1-3**;** and **C,** ATCC 10253.

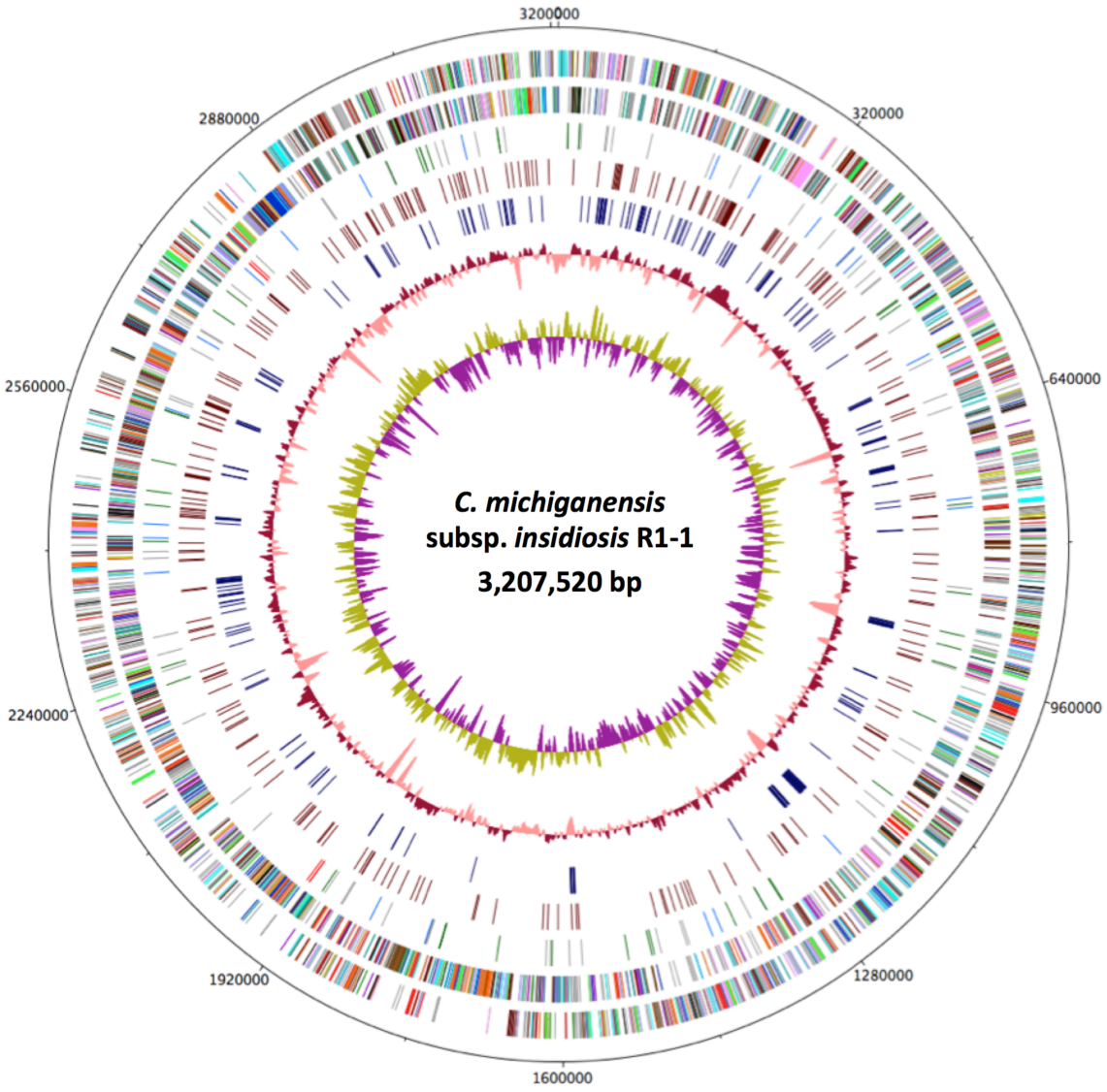


Figure 2. Circular plot of the chromosome of *Cmi* R1-1 with relevant features. From the outermost circle to the innermost circle: circle 1, nucleotide position (bp); circles 2 and 3, predicted protein coding sequences (CDS) in forward (circle 2) and reverse (circle 3) directions, with the colors indicating assigned Clusters of Orthologous Groups (COG) categories; circle 4, predicted tRNAs (in green), rRNAs (in red), insertion element IS*1122* (in blue) and pseudogenes (in grey); circle 5, predicted CDS containing signal peptides for secretion; circle 6, CDS that do not have orthologues in the genomes of *Cmm*, *Cms* and *Cmn*; circle 7, G+C content, above average shown in dark red and below average shown in light red (window size 5000 bp); and circle 8, G+C skew, higher values

shown in yellow, and lower values shown in purple (window size 5000 bp). Color keys for COG categories in circle 2 and 3: Red, Energy production and conversion; Green, Cell cycle control, cell division, chromosome partitioning; Orange, Amino acid transport and metabolism; Yellow, Nucleotide transport and metabolism; Maroon, Carbohydrate transport and metabolism; Pink, Coenzyme metabolism, secondary metabolites biosynthesis, transport and catabolism; Blue, Translation, ribosomal structure and biogenesis; Teal, Transcription; Azure, Replication, recombination and repair; Brown, Cell wall, membrane and envelope biogenesis; Light Blue, Posttranslational modification, protein turnover, chaperones; Light Green, Inorganic ion transport and metabolism; Black, Signal transduction mechanisms and defense mechanisms; Grey, Other. Data visualization was done with DNAPlotter (version 10.2) (Carver et al., 2009).

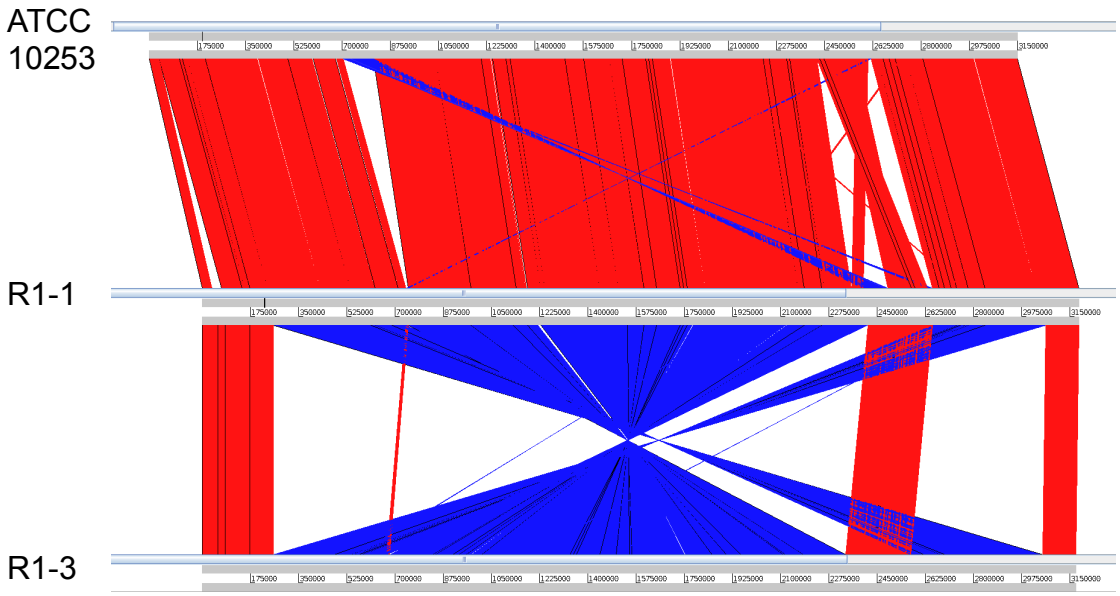


Figure 3. Alignment of chromosomes of *Clavibacter michiganensis* subsp. *insidiosus* strain ATCC 10253, R1-1 and R1-3. The start point of the chromosomal sequences of *Cmi* ATCC 10253 (top), R1-1 (middle) and R1-3 (bottom) are all set at the start codon of *dnaA*. The sequences of IS*1122* in R1-1 are removed to avoid alignments between copies of IS*1122*. Homologous regions larger than 1,000 bp in collinear or inverted direction are indicated by red or blue lines, respectively. Visualization of the alignment was performed with the Artemis Comparison Tool (ACT) (version 13.0.0) (Carver et al., 2005).

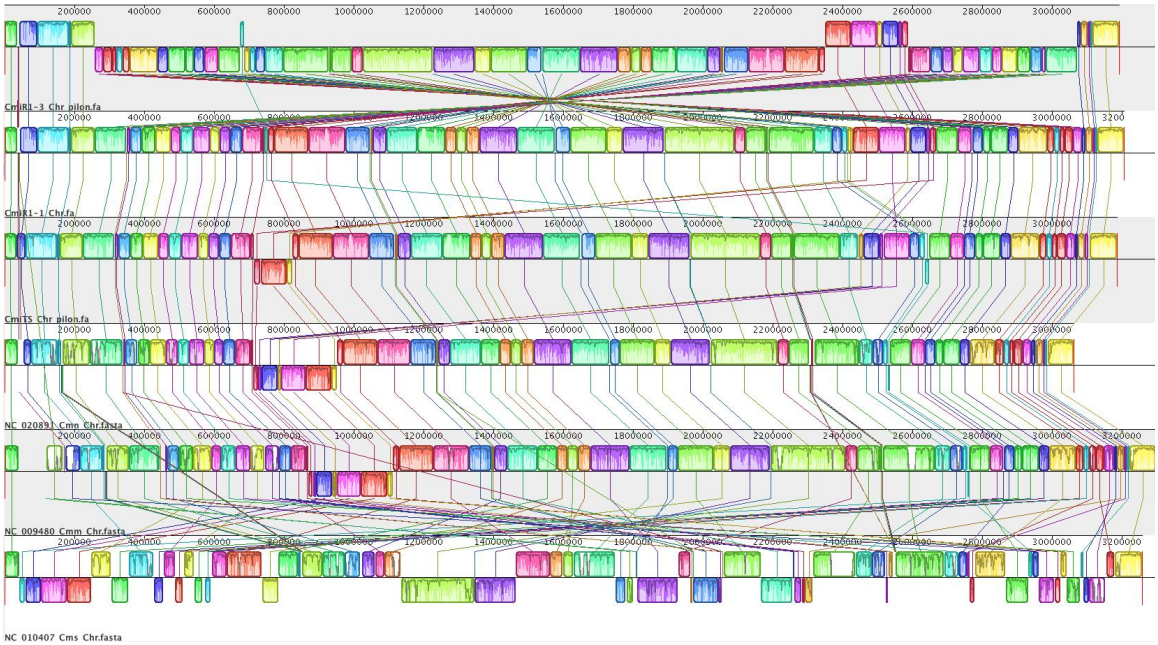
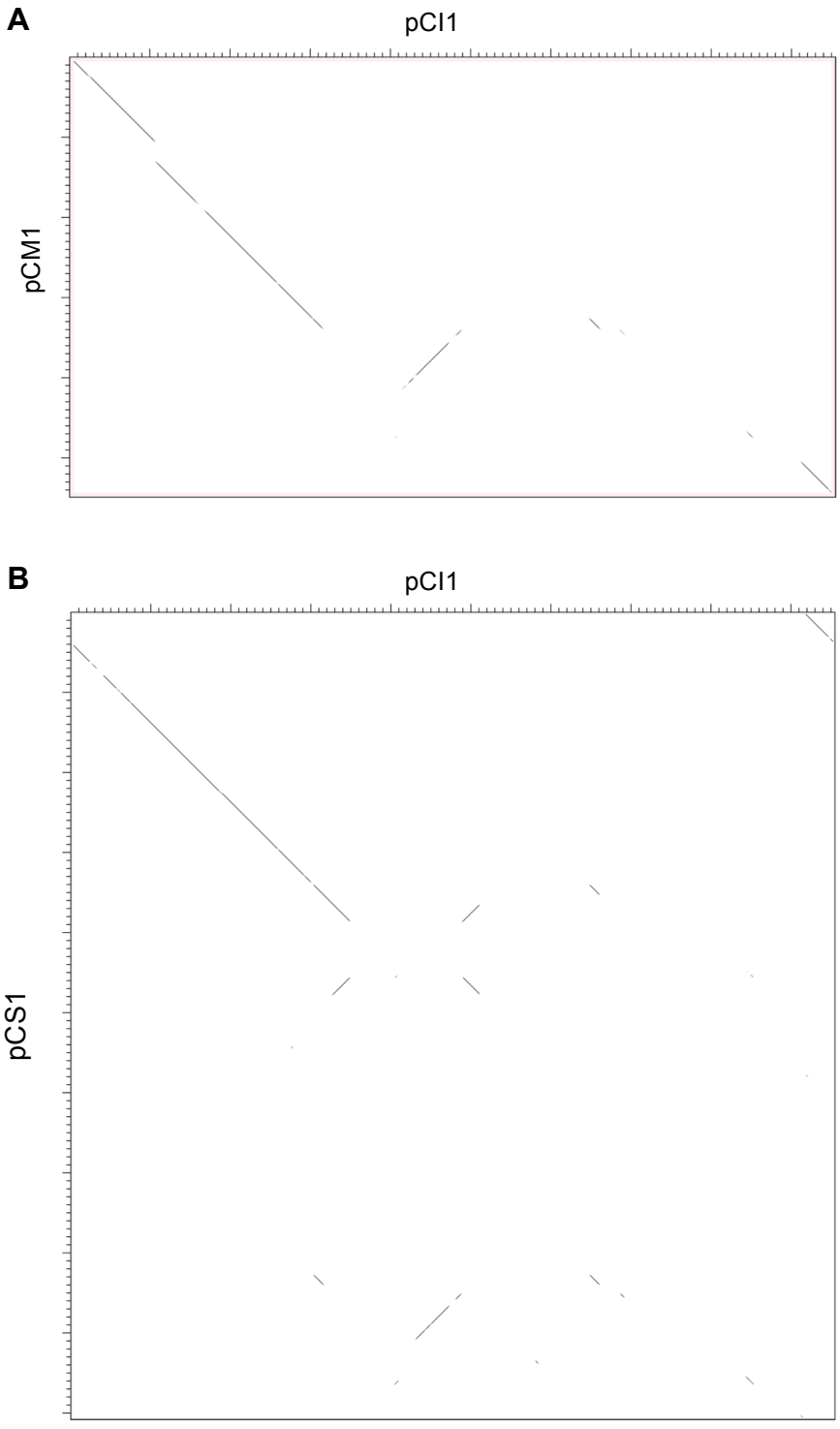


Figure 4. Whole chromosome alignment between the genomes of *Cmi*, *Cmm*, *Cmn* and *Cms*. The chromosomal sequences are shown as horizontal lines. From top to bottom: *Cmi* R1-3, *Cmi* R1-1 (as the reference genome), *Cmi* ATCC 10253, *Cmn* NCPPB 2581, *Cmm* NCPPB 382 and *Cms* ATCC 33113. Blocks with matching colors and connected with lines are considered locally collinear blocks (LCBs). LCBs appearing above each horizontal line are in the same orientation as the reference genome, and the ones below each horizontal line are in the reverse orientation relative to the reference genome. The height of the plots within each LCB depicts the level of sequence homology relative to the reference genome. Visualization of the alignment was done using Mauve (version 2.4.0) (Darling et al., 2004).


A
pCI1
pCM1
B
pCI1
pCS1

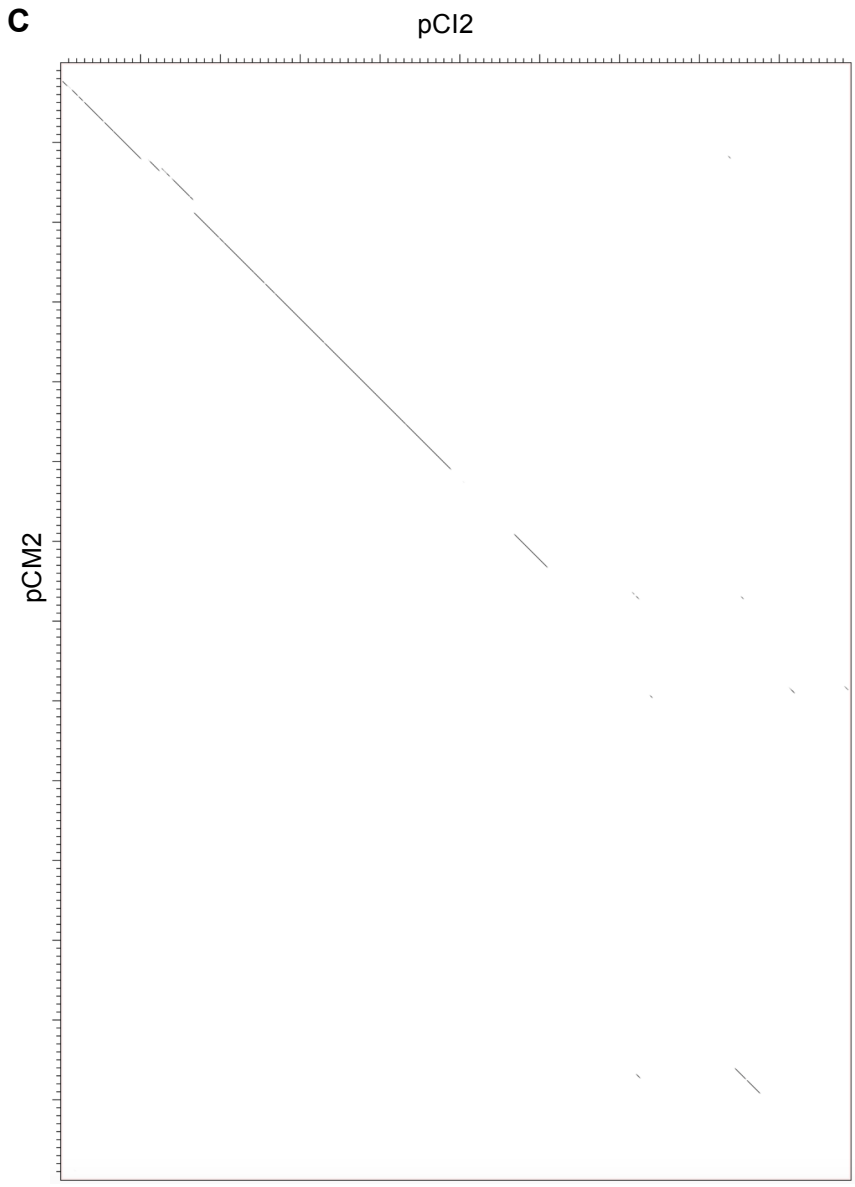


Figure 5. Dot-plot of alignments of plasmids in the genomes of *Cmi* R1-1, *Cmm* NCPPB 382 and *Cms* ATCC 33113. A, pCI1 (x-axis) vs. pCM1 (y-axis); **B,** pCI1 (x-axis) vs. pCS1 (y-axis); and **C,** pCI2 (x-axis) vs. pCM2 (y-axis). Visualization of the alignments was done using Dotter (version 4.34.3) (Sonnhammer and Durbin, 1995) with sliding window size set to 200. Major ticks on the axis represent 5,000 bp and minor ticks represent 500 bp.

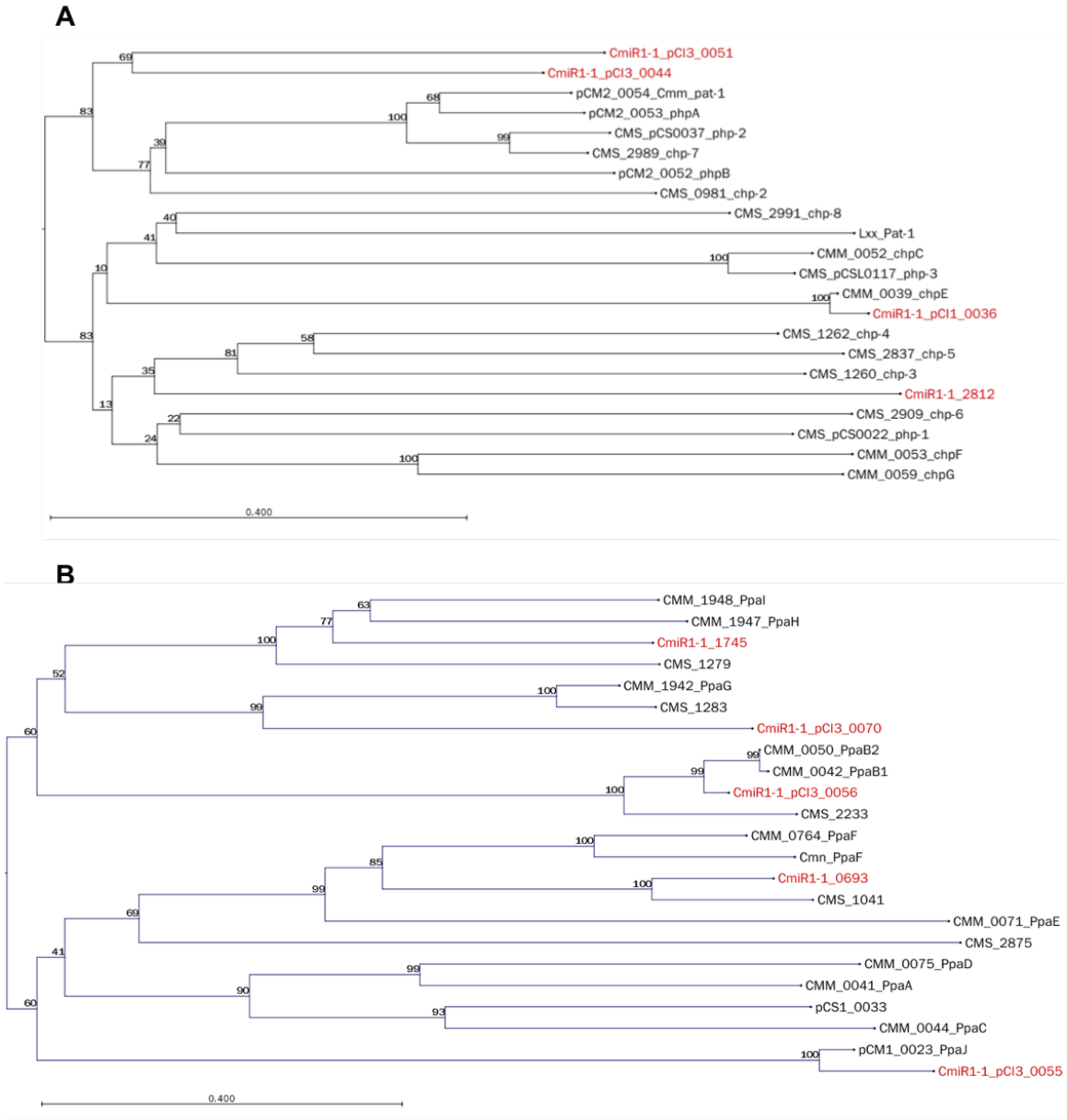


Figure 6. Maximum likelihood phylogenetic trees of the *pat-1* family and *ppa* family serine proteases. The full-length amino acid sequences of the gene products of **A,** the *pat-1* family and **B,** the *ppa* family were aligned with Multiple Sequence Comparison by Log- Expectation (MUSCLE). The phylogenetic trees were then created using the Neighbor Joining method with the WAG protein substitution model and 1,000 times bootstrap analysis in the CLC Main Workbench (version 7.6). The percentage bootstrap values are shown at the branching positions. The name of the *Cmi* gene products are highlighted in red.

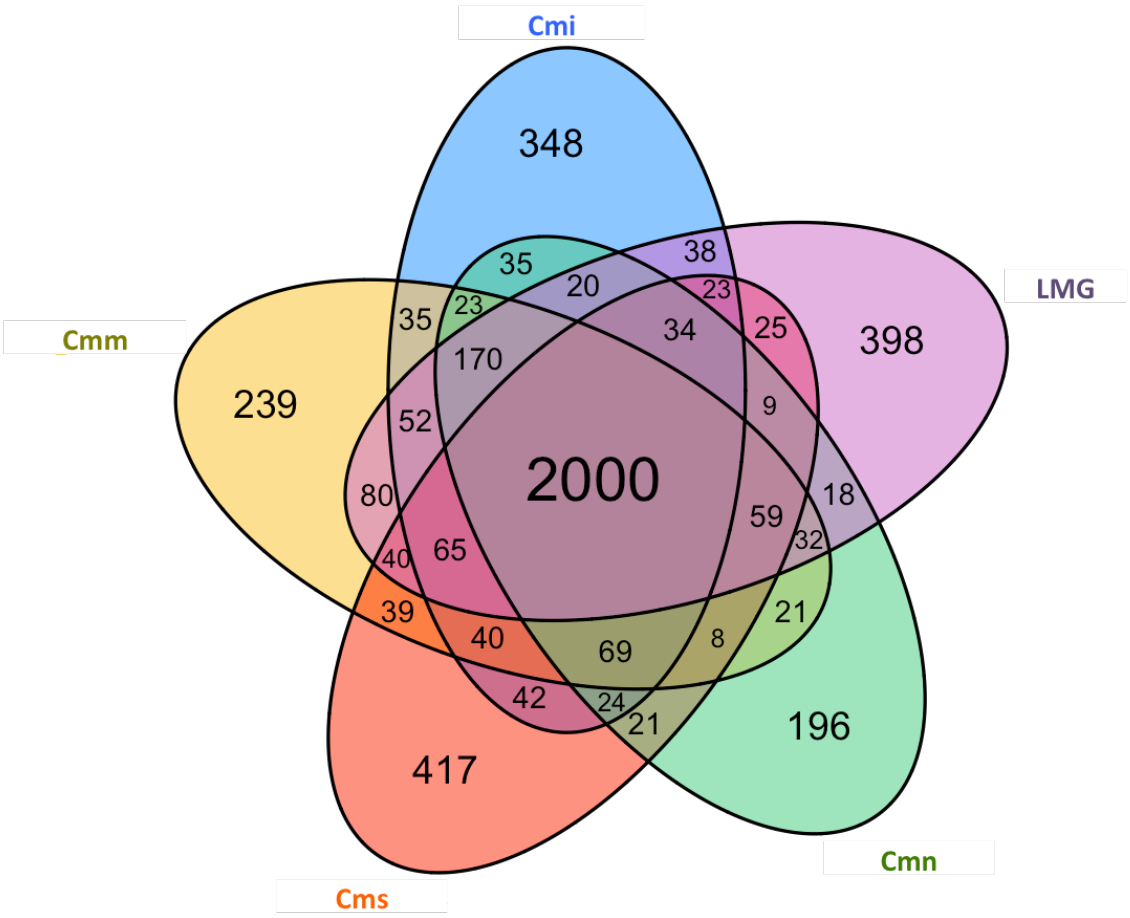


Figure 7. Venn diagram depicting the number of clusters of orthologous genes among *Cmi*, *Cmm*, *Cms*, *Cmn* and non-pathogenic *Clavibacter*. Numbers in the Venn diagram represent the number of clusters of orthologous genes analyzed with the software package GET_HOMOLOGUES (version 1.3) (Contreras-Moreira and Vinuesa, 2013). The input CDS from the five genomes were: **Cmi,** a pan-genome consisting of 3,168 CDS from *Cmi* R1-1, R1-3 and ATCC 10253; **Cmm,** CDS from *Cmm* NCPPB 382; **Cms,** CDS from *Cms* ATCC33113; **Cmn,** CDS from *Cmn* NCPPB2581; **LMG,** CDS from the nonpathogenic *Clavibacter* LMG 26808.

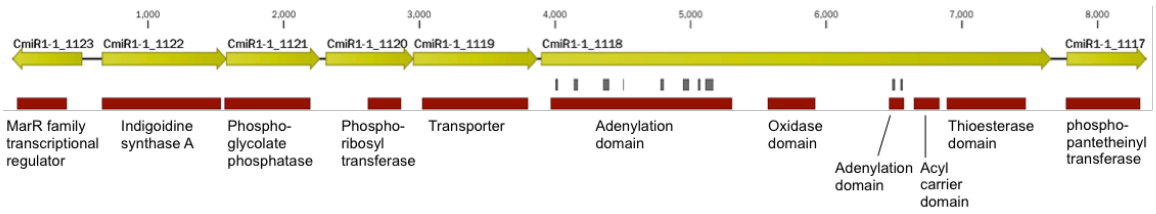


Figure 8. A cluster of genes in *C. michiganensis* subsp. *insidiosus* potentially involved in indigoidine biosynthesis. The predicted protein coding genes are shown in yellow and arrows indicate the directions of transcription. The identified domains or predicted functions are drawn below the genes at corresponding positions and are shown in red. The predicted AMP binding sites and conserved motifs were shown as grey bars above the adenylation domain. The scale on top is in base pair.

Chapter III. Putative serine protease effectors of *Clavibacter michiganensis* induce a hypersensitive response in the apoplast of *Nicotiana* species

Summary

Clavibacter michiganensis subsp. *michiganensis* (*Cmm*) and *C. michiganensis* subsp. *sepedonicus* (*Cms*) cause diseases on solanaceous crops. The genomes of both subspecies encode members of the *pat-1* family of putative serine proteases known to contribute to virulence on host plants and induction of hypersensitive responses (HR) on non-hosts. One gene of this family in *Cms*, *chp-7*, is required for triggering HR in *Nicotiana tabacum*. Here, further investigation revealed that mutation of the putative catalytic serine residue at position 232 to threonine abolished the HR induction activity of Chp-7, suggesting that enzymatic activity is required. Purified Chp-7 triggered an HR in *N. tabacum* leaves in the absence of the pathogen, indicating Chp-7 itself is the HR elicitor from *Cms*. Ectopic expression of *chp-7* constructs in *N. tabacum* leaves revealed that Chp-7 targeted to the apoplast triggered an HR, while cytoplasmic Chp-7 did not, indicating that Chp-7 induces the HR in the apoplast of *N. tabacum* leaves. Chp-7 also induced HR in *N. sylvestris*, a progenitor of *N. tabacum*, but not in other *Nicotiana* species tested. ChpG, a related protein from *Cmm*, also triggered HR in *N. tabacum* and *N. sylvestris*. Unlike Chp-7, ChpG also triggered HR in *N. clevelandii* and *N. glutinosa*.

Introduction

Well-adapted plant pathogens deploy effectors to suppress host defense responses, which is crucial for pathogen colonization and disease symptom development (Block et al., 2008). In Gram-negative plant pathogens, proteinaceous effectors are mainly delivered into the host cells via the type III secretion system (T3SS) (Alfano and Collmer, 2004). Some fungal and oomycete effectors secreted from pathogen haustoria are taken up by host plant cells mediated by an N-terminal RxLR-dEER motif in the effector proteins (Whisson et al., 2007; Dou et al., 2008). Certain pathogen effectors can be recognized directly or indirectly by plant resistance I proteins, leading to robust plant resistance often accompanied by a hypersensitive response (HR), a type of cell death within the infected tissue (Caplan et al., 2008). Recognition may be direct, as exemplified by recognition of AVR-Pita from *Magnaporthe grisea* by the rice R protein Pi-ta (Jia et al., 2000). It may also be indirect, conceptualized as the "Guard Hypothesis" in which a plant target (the guardee) is monitored by an R protein (the guard) for modifications by a pathogen effector. Conformational changes in or loss of the guardee protein activate the guard R protein (Dangl and Jones, 2001). For example, the Arabidopsis R protein RPS5 recognizes the *Pseudomonas syringae* type III effector AvrPphB, which is a cysteine protease, by monitoring the associated guardee host protein PBS1, which is cleaved by AvrPphB (Shao et al., 2003; DeYoung et al., 2012). Recognition of effectors by R proteins can occur in intra- or inter-cellular spaces, though examples of the latter are relatively rare (Martin et al., 2003). Perhaps the best-studied effector-R protein pairs that occur in the apoplast consist of effectors secreted by the fungal pathogen *Cladosporium fulvum* and tomato Cf receptor-like proteins (RLPs), such as the fungal apoplastic effector Avr2 and tomato Cf-2 (Rooney *et al.* 2005). Avr2 inhibits the host apoplastic protease Rcr3 leading to a conformational change that is in turn recognized by the R protein Cf-2 (van't Klooster *et al.* 2011).

Clavibacter michiganensis is a diverse species of Gram-positive bacteria in the Microbacteriaceae. Most characterized isolates are plant pathogens, but nonpathogenic isolates also occur (Zaluga *et al.* 2013). Currently there are six subspecies identified within *C. michiganensis* based on the host plants they colonize (Eichenlaub and Gartemann, 2011; Gonzalez and Trapiello, 2014). Although they cause diseases on

various major crop species, including tomato, potato, maize, alfalfa, wheat and bean, each of the *C. michiganensis* subspecies only infects a narrow range of plant hosts. For example, *C. michiganensis* subsp. *michiganensis* (*Cmm*) is the causal agent of bacterial canker disease on tomato (*Solanum lycopersicum*), and can infect the model plant *Nicotiana benthamiana* (Balaji et al., 2011). *C. michiganensis* subsp. *sepedonicus* (*Cms*) causes bacterial ring rot disease on cultivated potato (*Solanum tuberosum*), and causes disease symptoms relatively quickly in inoculated eggplant and tomato, often used as artificial hosts for virulence assays (Larson 1944).

In *C. michiganensis*, full virulence on host plants is associated with several virulence factors. Symptom development caused by *Cmm* in tomato requires the two native plasmids, pCM1 and pCM2. pCM1 carries *celA*, which encodes a β-1,4-endocellulase, and is required for full virulence on tomato (Meletzus et al., 1993). *Cms* contains a gene homologous to *celA*, which is also required for disease symptom development on potato (Laine et al., 2000). pCM2 carries *pat-1* (pCM2_0054), which encodes a putative serine protease of subfamily S1A, and is also required for full virulence (Dreier et al., 1997). Additional genes encoding putative serine proteases homologous to *pat-1* have been identified in the *Cmm* genome, named *chromosomal homology of pat-1* (*chp*) and *plasmidal homology of pat-1* (*php*) genes (Burger et al., 2005; Gartemann et al., 2008). All the *chp* genes are clustered in a 129 kb pathogenicity island in the *Cmm* chromosome with lower than genome average %G+C content, suggesting they were acquired via horizontal gene transfer (Gartemann et al., 2008). Alignment of the amino acid sequences encoded by *pat-1* homologs revealed the conserved trypsin family serine protease motifs that contain the predicted catalytic histidine and serine residues (Stork et al., 2008). A loss-of-function mutation in one of the *chp* genes, *chpC* (CMM_0052), compromised colonization of *Cmm* and symptom development on tomato (Stork et al., 2008), and the expression of *chpC* is dependent on the presence of both plasmids in *Cmm* (Chalupowicz et al., 2010). In *Cms*, loci encoding homologs of *pat-1* are also present on the chromosome and plasmids, though they are not clustered within a pathogenicity island, presumably due to extensive genomic rearrangement events (Bentley et al., 2008). One close homolog of *pat-1*, named *chp-7* (CMS_2989), has been identified as an important virulence factor of *Cms* required for

full symptom development in infected potato and eggplant but not for colonization (Nissinen et al., 2009). Like Pat-1, Chp-7 contains a predicted catalytic histidine at position 97 within a TAKHC motif and the predicted nucleophilic serine at position 232 within a GDSGG motif (Nissinen et al., 2009). Though no proteolytic activity has been demonstrated for any of the *pat-1* family serine proteases, the serine protease activity of Pat-1 is thought to be involved in virulence of *Cmm* in tomato, as exchanging the putative catalytic serine for threonine in Pat-1 rendered it unable to restore the virulence of the plasmid-free strain CMM100 (Burger et al., 2005).

Both *Cmm* and *Cms* are known to secrete one or multiple proteinaceous HR elicitors (Alarcon, 1998 #3341;Nissinen, 1997 #3229). It has been reported that a *chpG* (CMM_0059) mutant of *Cmm* and a *chp-7* mutant of *Cms* lost the HR eliciting ability in *Mirabilis jalapa* and *N. tabacum*, respectively, indicating that ChpG and Chp-7 are required for HR (Stork et al., 2008). However, it is not known if these proteins are the direct HR elicitors.

Gram-positive plant pathogenic bacteria do not have any secretion system known to be functionally equivalent to the Gram-negative T3SS. Most of the studied and suspected *C. michiganensis* virulence factors carry a signal peptide for secretion, suggesting they are secreted. This idea is supported by a recent proteomics study using the *Cmm*-tomato pathosystem, in which several of the *pat-1* family serine proteases, *ppa* family serine proteases, CelA, and other hydrolases, were detected in both culture supernatants of bacteria grown in tomato xylem sap-supplemented minimal medium and the xylem sap from infected tomato plants (Savidor et al., 2012). Chp-7 contains a signal peptide similar to the one in Pat-1 and is predicted to be a secreted protein (Nissinen et al., 2009). It is not known whether any of the *Cmm* or *Cms* secreted proteins enter plant cells.

In this study, we investigated the Chp-7-triggered HR in *N. tabacum*. We found that the HR-inducing ability requires the predicted catalytic serine residue of Chp-7. Purified Chp-7 triggered HR, indicating that Chp-7 is sufficient for triggering an HR and is the recognized elicitor. The recognition of Chp-7 by *N. tabacum* requires that *Chp-7* is present in the apoplast. Both Chp-7 and ChpG are recognized in *N. sylvestris*, one of the two progenitor species of *N. tabacum*. ChpG, but not Chp-7, is recognized in *N.*

clevelandii and *N. glutinosa* as well, indicating that the two proteins do not have identical activities.

Results

A point mutation in the predicted catalytic serine abolished the HR-inducing ability of Chp-7

Alignment of the amino acid sequences of all the *pat-1* homologs from *Cmm* and *Cms* revealed the highly conserved serine protease motif GDSG(G/A/C)(P/F/A)V containing the predicted catalytic serine residue (Figure 1A). To test whether the conserved serine residue at position 232 of Chp-7 is required for inducing HR in *N. tabacum* leaves, a complementation assay was carried out by introducing either the wild type *chp-7* sequence or *chp-7ST* containing a Ser232Thr point mutation (Figure 1A) into CmsC7X20, a *chp-7* knockout mutant derivative of *Cms* ATCC33113, previously created by inserting the chloramphenicol resistance cassette *cmx* into the coding region of *chp-7* via homologous recombination (Syverson 2011). Both constructs were expressed from the native promoter of *chp-7* and were fused to a C-terminal FLAG epitope tag. The constructs were cloned into the *E. coli-Clavibacter* shuttle vector, pHN216 (Laine et al., 1996), resulting in pHNC7-F and pHNC7ST-F. The transformed CmsC7X20 strains were then assayed for the ability to induce HR in infiltrated *N. tabacum* leaves. Chp-7-FLAG restored the HR-inducing ability in CmsC7X20, while Chp-7ST-FLAG did not (Figure 1B). Culture supernatants from CmsC7X20_216, containing the pHN216 empty vector, CmsC7X20_C7-F, expressing Chp-7-FLAG, and CmsC7X20_C7ST-F, expressing Chp-7ST-FLAG, were collected. Immunoblotting showed that both Chp-7-FLAG and Chp-7ST-FLAG were secreted from CmsC7X20 and accumulated to similar levels in culture supernatants (Figure 1C). These results indicated that the predicted catalytic serine residue in Chp-7 is required for its HR-inducing ability in *N. tabacum*.

Chp-7 triggers an HR in *N. tabacum* leaves in the absence of pathogen

Because Chp-7 is a secreted protein, it may be sufficient to trigger HR in *N. tabacum* leaves. To test this hypothesis, an epitope-tagged construct, *chp-7-6×His* driven by an IPTG-inducible promoter, was synthesized and introduced into *E. coli* for

expression. Tandem histidine-tagged proteins from cell lysates of *E. coli* cultures induced with IPTG or uninduced were purified with cobalt-based affinity resin. The eluted proteins were analyzed by gel electrophoresis and immunoblotting. Chp-7-6×His was detected as a band at the predicted mass of 28 kDa, present in the IPTG-induced culture, but not in uninduced cultures or cultures lacking the recombinant construct (Figure 2). Several contaminating *E. coli* proteins were present in eluates from all cultures. The eluted proteins were infiltrated into *N. tabacum* leaves following a dilution series. HR was observed in the areas infiltrated with the diluted eluate from the IPTG-induced *E. coli* culture, but not in areas infiltrated with eluate from the uninduced culture even at undiluted level (Figure 2). The intensity of the cell death decreased as the Chp-7-6×His concentration was reduced. This showed that Chp-7 itself is the HR elicitor and can trigger HR in *N. tabacum* leaves in the absence of *Cms*.

Chp-7 is not translocated into plant cells

The genome of *C. michiganensis* does not contain genes for a secretion system known to deliver bacterial proteins into plant cells. That does not rule out the possibility that *C. michiganensis* is capable of translocating proteins into host cells through an as yet unknown mechanism. We were interested to determine whether Chp-7 is translocated into plant cells following secretion by *Cms*. To measure Chp-7 translocation, we used an adenylate cyclase (Cya) fusion assay that has been widely used to detect translocation of pathogen effectors into plant cells (Casper-Lindley et al., 2002; Schechter et al., 2004). Candidate translocated proteins are fused to the catalytic domain of adenylate cyclase of the cyclolysin from *Bordetella pertussis*. Cya is activated in a calmodulin-dependent manner and catalyzes the conversion of ATP into 3',5'-cyclic AMP (Sory and Cornelis, 1994). Detection of cAMP in infiltrated leaves serves as the measurement of candidate-Cya translocation. Accordingly, we constructed a *chp-7-cya* fusion. The activity of the Chp-7-Cya fusion protein was first tested in *N. tabacum* leaves by infiltrating the leaves with CmsC7X20_C7Cya, expressing the Chp-7-Cya fusion. The Chp-7-Cya fusion protein triggered HR to a similar extent as the wild type Chp-7, indicating Chp-7-Cya is active for HR induction (Figure 3A). Because the onset of HR might interfere with cAMP production, the measurement of cAMP was carried out using infiltrated *N. benthamiana*

leaves, as Chp-7 does not induce HR in this species. An Effector-to-Host Analyzer (EtHAn) strain (Thomas et al., 2009) carrying the Cya construct with a T3SS signal (T3sp-Cya) was used as a positive control. Leaf samples from infiltrated areas were collected at 0, 10 and 24 hours post-inoculation (hpi). The leaf samples infiltrated with EtHAn:Cya showed approximately a thousand-fold increase in cAMP concentration at 10 and 24 hpi compared to 0 hpi. The leaf samples infiltrated with CmsC7X20_C7Cya did not show increased cAMP production during the same time course (Figure 3B), which suggests Chp-7-Cya is not translocated into plant cells following secretion from *Cms*.

Transient expression of *chp-7* secreted into the apoplast triggers HR in *N. tabacum* leaves

To determine in which subcellular plant compartment Chp-7 acts to trigger an HR, three different *chp-7* constructs with or without a plant secretion signal were made for transient expression in *N. tabacum* (Figure 4A). The *chp-7*$_{Apop}$ and *chp-7ST*$_{Apop}$ constructs contain an N-terminal plant signal peptide from tobacco PR1a (amino acids 1-31) known to target protein accumulation into the apoplast (Wulff *et al.*, 2004; van Esse *et al.*, 2006). The *chp-7*$_{Apop}$ construct has the wild type *chp-7* sequence, while *chp-7ST*$_{Apop}$ contains a $Ser_{232}Thr$ point mutation in the sequence. The *chp-7*$_{Cyto}$ construct differs from *chp-7*$_{Apop}$ only in that it omits the signal peptide sequence from *PR1a* and should direct Chp-7 accumulation to the cytoplasm of plant cells. All three constructs contain the intron *PIV2* from the potato *ST-LS1* gene, in order to prevent expression in *Agrobacterium tumefaciens*, and a C-terminal HA-PreScission-Biotin (HPB) tandem epitope tag (Qi and Katagiri, 2009). The recombinant constructs were inserted into binary vector pMDC32, expressed from a cauliflower mosaic virus (CaMV) *35S* promoter for expression in *N. tabacum*. Following infiltration of *N. tabacum* leaves with *A. tumefaciens* strains carrying the recombinant constructs, apoplastic Chp-7 triggered HR at 2 days post-inoculation (dpi), while neither secreted Chp-7^{S232T} nor cytoplasmic Chp-7 did (Figure 4B). Amounts of all three versions of recombinant Chp-7 proteins were visualized by 57ictor-detection with anti-HA antibody. Total proteins were extracted from the infiltrated leaves at 12 hpi, just before the HR triggered by apoplastic Chp-7 became visible. Proteins produced by *chp-7ST*$_{Apop}$ or *chp-7*$_{Cyto}$ accumulated to a much higher level compared to the amount of

apoplastic Chp-7 (Figure 4C). The expression of *chp-7*$_{Apop}$ is likely suppressed by the onset of HR, explaining its low accumulation. Biotinylation occurs predominantly in the cytosol, and to a lesser extent in mitochondria and chloroplasts (Chapman-Smith and Cronan, 1999). Proteins exported to the apoplast via the endoplasmic reticulum (ER)-Golgi mediated secretory routes are thus not biotinylated. Cytoplasmic Chp-7 was biotinylated while apoplastic Chp-7ST was not (Figure 4D), which is consistent with their localization in the cytoplasm or apoplast, respectively. Therefore, the transient expression of constructs including the *PR1a* signal peptide did deliver Chp-7 proteins into the apoplast, while the transient expression of *chp-7*$_{Cyto}$ led to protein accumulation in the cytoplasm. Evidently, Chp-7 is only effective in inducing HR when it is present in the apoplast. Taken together, these experiments strongly suggest Chp-7 is recognized in the apoplast of *N. tabacum* leaves.

ChpG from *C. michiganensis* subsp. *michiganensis* is also recognized by *N. tabacum* in the apoplast

Cmm secretes one or multiple proteins that can trigger HR in *Mirabilis jalapa* and *N. tabacum* (Alarcon et al., 1998). So far only *chpG* was reported to be required for *Cmm* to trigger an HR in *M. jalapa* (Stork *et al.* 2008), however, the identity of the elicitor(s) in *N. tabacum* remains unknown. The genome of *Cmm* contains seven intact *pat-1* homolog genes (Gartemann *et al*. 2008). Among these, the pCM2 plasmid-encoded *pat-1*, *phpA* and *phpB* have the highest amino acid sequence similarity to *chp-7*. Therefore, we tested the idea that one or more of these genes functions in a similar way as *chp-7*. The Cm15-2.0 sm strain (*Cmm* wild-type), the CMM100 strain (*Cmm ΔpCM1ΔpCM2*) and the CMM101chpGβ (*Cmm ΔchpGΔpCM2*) strain were infiltrated in *N. tabacum* leaves, and responses were examined at 2 dpi. While HR occurred in the area infiltrated with Cm15-2.0 sm or CMM100, it was not observed in the leaf area infiltrated with CMM101chpGβ, while some chlorosis developed (Figure 5A). Although CMM101chpGβ lacks *chpG*, *pat-1*, *phpA* and *phpB*, CMM100, lacking *pat-1*, *phpA* and *phpB* but having *chpG*, is capable of triggering HR, indicating only *chpG* is required for triggering an HR in *N. tabacum*, while *pat-1*, *phpA* and *phpB*, encoded on pCM2, are not. Consistently, full-length *pat-1*, *phpA* or *phpB*, expressed from the *chp-7* promoter all failed to restore the HR-triggering

phenotype of CmsC7X20 in a complementation assay (Figure 7). To test if apoplastic accumulation of ChpG is sufficient to trigger an HR in *N. tabacum*, a recombinant construct encoding the tobacco PR1a signal peptide fused to ChpG and expressed from the CaMV *35S* promoter was delivered into *N. tabacum* leaves through *Agrobacterium*-mediated transformation. A similar construct of ChpC was made and expressed. HR cell death was observed following transient expression of the *PR1a-ChpG* construct, but not in the leaf area transiently expressing the *PR1a-ChpC* fusion protein. These results show that ChpG is the HR elicitor of *Cmm* in *N. tabacum*, and that ChpG, like Chp-7, is recognized in the apoplast.

The ability to recognize Chp-7 in *Nicotiana* species arises from the Sylvestres section

N. tabacum is an allotetraploid plant whose S and T genomes are derived from two wild progenitors, *N. sylvestris* and *N. tomentosiformis*, respectively (Kenton et al., 1993). To estimate the origin of the gene(s) required for recognizing Chp-7 and ChpG in *N. tabacum*, wild-type and mutant strains of *Cmm* and *Cms* were infiltrated into leaves of *N. sylvestris* and *N. tomentosiformis*. *N. sylvestris* behaved like *N. tabacum*, as HR was triggered by ATCC33113 (*Cms* wild-type), Cm15-2.0 sm (*Cmm* wild-type) and CMM100 (*Cmm ΔpCM1ΔpCM2*) but not by CmsC7X20 (*Cms Δchp-7*) or CMM101chpGβ (*Cmm ΔchpGΔpCM2*). Transient expression of apoplastic Chp-7 and ChpG both triggered HR in *N. sylvestris,* demonstrating that these proteins are sufficient for triggering HR in this species as well. In contrast, none of the strains or transiently expressed proteins triggered HR in *N. tomentosiformis* (Figure 5B and 5C). Interestingly, all the *Cmm* strains induced some chlorosis in both *N. sylvestris* and *N. tomentosiformis*, while *Cms* strains did not. These results show that the ability of *N. tabacum* to respond to Chp-7 and ChpG is encoded in the S genome.

To test whether the ability to recognize Chp-7 and ChpG is a common trait among *Nicotiana* species, strains of *Cms* and *Cmm* were infiltrated into leaves of *N. glutinosa*, *N. 59ictoria59*, *N. rustica*, *N. clevelandii*, *N. benthamiana* and *N. glauca*, each representing a section in the *Nicotiana* phylogeny. *N. clevelandii* and *N. glutinosa* both developed a *chpG*-dependent HR in response to *Cmm*, though in *N. glutinosa* the ChpG-triggered HR took more than 2 days to become visible (Figure 6). In both species, transient expression

of the *PR1a-ChpG* recombinant construct was sufficient to trigger HR. However, neither of the two species recognized *Cms* or Chp-7. The foliar responses at 3 dpi of all the *Nicotiana* species tested are summarized in Table 2. As some species recognize ChpG but not Chp-7, the activities of these proteins must be different, even though they induce similar responses in *N. tabacum*.

Discussion

Extensive research on Gram-negative phytopathogenic bacteria, such as *Pseudomonas syringae*, has provided insights into pathogen virulence strategies and host defense responses. However, much less is known about their Gram-positive plant pathogenic counterparts in the family Microbacteriaceae, such as *Clavibacter michiganensis.* Gram-positive bacteria lack outer membranes, and most have a thick cell wall comprised of peptidoglycan. These distinct structural differences relative to Gram-negative bacteria necessitated evolution of different strategies for interacting with plant hosts. Lacking a secretion system equivalent to Gram-negative T3SS and T4SS makes the apoplastic space the primary interface for the interactions between *C. michiganensis* and plants. In this study, we showed that effector proteins secreted by *Cms* and *Cmm* are recognized in the apoplast in *N. tabacum*.

Gene-for-gene type resistance is mediated by plant R proteins and specifically recognized pathogen effectors. The only group of R proteins known to recognize extracellular effectors are the RLP class, including tomato Cf proteins and Ve proteins, which contain extracellular LRR (eLRR) domains (van Ooijen et al., 2007). The R protein Cf-2 initiates an HR upon detecting conformational changes in the extracellular cysteine protease Rcr3 induced by binding of *C. fulvum* Avr2 as an uncompetitive inhibitor (van't Klooster *et al.* 2011). The recognition of Chp-7 and ChpG in the apoplast may be mediated by an R protein(s) from *N. tabacum* that contains an eLRR domain. As the substrates of Chp-7 and ChpG remain unidentified at this point, it is equally possible that the recognition occurs by Chp-7 or ChpG interacting with a guardee protein(s) in the apoplast monitored by an R protein or by direct interaction with the R protein(s) as its substrate, though to our knowledge there is no known example for the latter scenario.

The *Cmm* genome has ten *pat-1* homologs (among them *chpA*, *chpB* and *chpD* are pseudogenes), and the *Cms* genome encodes eleven *pat-1* homologs (only *chp-1* is a pseudogene) (Bentley et al., 2008; Gartemann et al., 2008). Single mutations in *pat-1* homolog genes including *pat-1*, *chpC* and *chp-7*, resulted in compromised symptom development or *in planta* colonization (Dreier et al., 1997; Stork et al., 2008; Nissinen et al., 2009), indicating the *pat-1* family of putative serine proteases play crucial and non-redundant roles for pathogen virulence. In addition, the observation that *N. clevelandii* and *N. glutinosa* can recognize ChpG but not Chp-7 indicates different plant proteins are involved. The lack of redundancy suggests that they function independently, perhaps by interacting with different host targets.

Interestingly, we observed chlorosis symptoms in the leaves of multiple *Nicotiana* species infiltrated with *Cmm*, while *Cms* induced chlorosis only in *N. rustica* (Table 2). The chlorosis likely occurred independently of ChpG, as the intensity of chlorosis induced by CMM101chpGβ was indistinguishable from those by other *Cmm* strains (Figure 5C). In the *Nicotiana* species that can recognize ChpG, such as *N. tabacum*, this *Cmm*-induced chlorosis symptom was likely masked by HR. The cause of the chlorosis and whether it is associated with virulence of *Cmm* remain to be elucidated in further investigations.

The expression of most of the studied virulence factors secreted by *C. michiganensis*, including Pat-1, CelA and other cell wall degrading enzymes, are inducible upon entering host tissues or upon addition of xylem sap into culture media (Chalupowicz et al., 2010). Moreover, the induced expression of virulence genes located on plasmids or within the pathogenicity island on the chromosome, called the PAI region, is interdependent. Deletion of either plasmid of *Cmm* results in dramatic reductions in expression of genes within the PAI region, and deletion of the PAI region also led to reduced induction of *pat-1* and *celA* expression (Chalupowicz et al., 2010). However, *chp-7* is not inducible during infection of potato leaves (Holtsmark et al., 2008), and the protein extract precipitated from culture supernatant of *Cms* triggered HR (Nissinen et al., 1997). In this study, Chp-7-FLAG was detected in the supernatant of YGM broth cultures (Figure 1C), suggesting it is constitutively expressed in culture in rich defined medium, which is consistent with the previous observations.

Pat-1 and Chp-7 both contain an LPGSG motif in the C-terminal region, which resembles the LPxTG sortase-processing motif for sortase-mediated covalent attachment to the cell wall of Gram-positive bacteria (Burger et al., 2005; Nissinen et al., 2009). However, the detection of Chp-7-FLAG in the culture supernatant by immunoblot with monoclonal anti-FLAG antibody suggests the C-terminal region downstream of the LPGSG motif was not cleaved by the *Cms* sortase. Therefore, the LPGSG motif in Chp-7 is either not a *bona fide* sortase-processing motif, or at least a subpool of Chp-7 proteins are not anchored to the *Cms* cell wall and HR induction can be independent of cell wall anchoring of Chp-7.

The substrates for the putative serine proteases involved in pathogenicity or induction of HR by *C. michiganensis* remain unknown. There is no detectable proteolytic activity in *Cmm* culture supernatants assayed with various generally used substrates (Burger et al., 2005) or using purified Chp-7-6×His from *E. coli* assayed with azocoll (this study, data not shown). Although the conserved motifs in the sequences of *pat-1* family genes resemble trypsin family serine proteases, they may also be some other kind of hydrolase, with a non-protein substrate. If they are proteases, they may have strict substrate specificity. Furthermore, though the amino acid sequences of Chp-7 and Pat-1 share 78% identity, Pat-1 cannot restore the HR induction in the *chp-7* mutant strain (Figure 7), suggesting the variable residues dictate interactions with their specific substrates. Nevertheless, the point mutation replacing the serine at position 232 with threonine in Chp-7 abolished its HR inducing ability in *N. tabacum* (Figure 1B and Figure 4B), indicating the serine residue predicted to function as a nucleophile is required in this context and the behavior of Chp-7 is recognized by *N. tabacum*. The substrate of Chp-7 is likely of plant origin, as the purified Chp-7-6×His is able to trigger HR in the absence of *Cms* (Figure 2).

We demonstrated that Chp-7 is sufficient to trigger an HR in the apoplast of *N. tabacum* leaves. By analogy to the Cf system in tomato, we hypothesize that Chp-7 cleaves a plant-derived apoplastic target, leading to conformational changes in the target, which in turn activate an eLRR domain-containing R protein, and HR ensues. Investigation of the virulence functions of Chp-7 and identification of Chp-7 substrate(s) in *N. tabacum* and the cognate R protein would allow this hypothesis to be tested.

Materials and Methods

Bacterial strains and growth conditions

The bacterial strains used in this study are listed in Table 1. *C. michiganensis* wild type strain was cultured on YGM medium (De Boer and Copeman 1980), CmsC7X20 was cultured on YGM medium supplemented with 10 μg/ml chloramphenicol, and CmsC7X20 strains transformed with constructs in the pHN216 vector were cultured on YGM medium supplemented with 10 μg/ml chloramphenicol and 25 μg/ml kanamycin. The *chp-7* knock out mutant CmsC7X20 was generously provided by R.L. Syverson. It was created by disrupting the *chp-7* coding region in the genome of *Cms* with a chloramphenicol resistance cassette, *cmx*, through homologous recombination with a suicide vector pRLSTC7X2 (Syverson 2011). CMM100 (Meletzus and Eichenlaub 1993) and CMM101chpGβ (Stork *et al.* 2008) have been described. All *C. michiganensis* cultures were started from glycerol stocks at -80 °C and grown at 25.5 °C for 4 days before use. *Escherichia coli* BL21 (DE3) pLysS (Invitrogen, Carlsbad, CA) and its derivatives were grown on Luria-Bertani (LB) medium at 37 °C and supplemented with 34 μg/ml chloramphenicol and 50 μg/ml kanamycin as necessary. *Agrobacterium tumefaciens* GV3101 (Koncz and Schell, 1986) and its derivatives were grown on LB medium at 28°C and supplemented with 20 μg/ml gentamycin and 25 μg/ml kanamycin as necessary.

Cloning and site-specific mutagenesis of *chp-7*

For complementation studies, CmsC7X20 was transformed with pHNC7 or pHNC7ST created by inserting wild type *chp-7* or *chp-7*^{S232T} into shuttle vector pHN216 (Laine *et al.* 1996). The *chp-7*^{S232T} derivative was made by PCR using KOD Hot Start DNA polymerase (Novagen, CA) using pRLSTC7 (Syverson 2011) as template with primers C7_S232T_F and C7_S232T_R (Table 3) followed by self-ligation with T4 DNA ligase (New England Biolabs, MA), resulting in pRLSTC7ST. The EcoRI fragment containing *chp-7*^{S232T} was ligated into pHN216, resulting in pHNC7ST. C-terminal FLAG epitope tags were inserted into pHNC7 and pHNC7ST by PCR with primers C7_OlpHN_F, C7_FLAG_R and FLAG_OlpHN_R, followed by ligation using the

Gibson assembly cloning kit (New England Biolabs, MA) into pHN216 cut open with EcoRI, resulting in pHNC7-F and pHNC7ST-F. Both plasmids were transformed into competent cells of CmsC7X20 by electroporation at 1.5 kV in cuvettes with a 1 mm gap width as previously described (Nissinen *et al*. 2009). Colonies were selected on YGM medium supplemented with 10 μg/ml chloramphenicol and 25 μg/ml kanamycin.

The DNA sequence of *chp-7* was commercially codon-optimized for expression in *E. coli*. The optimized sequence, *chp-7E,* was cloned into expression vector pJExpress401 (DNA2.0, CA), resulting in pJEC7H. The plasmid was transformed into *E. coli* strain BL21 (DE3) pLysS (Invitrogen, CA) using the heat-shock method according to the manufacturer's protocol. Colonies were selected on LB medium supplemented with chloramphenicol and kanamycin.

The entry clones for transient expression experiments were created by TOPO TA cloning of the *chp-7* or *chp-7*^{S232T} coding sequences, omitting the native signal peptide region, using the primers Chp7v35_F and C7+279R, into pCR8/GW/TOPO vector (Invitrogen, CA), resulting in pCR8_C7 and pCR8_C7ST, respectively. All the *chp-7* constructs used for transient expression in plants contain the second intron *PIV2* from the potato *ST-LS1* gene (Vancanneyt et al., 1990), which was cloned from pAD1339 (Ke et al., 2001) using primers OLC7KpnI_PIV2_F and OLC7KpnI_PIV2_R. The resulting fragment contains sequences overlapping ones flanking the unique KpnI site in the wild type *chp-7* sequence. The *PIV2* fragment was ligated using the Gibson assembly cloning kit (New England Biolabs, MA) into pCR8_C7 and pCR8_C7ST cut open with KpnI. The *chp-7* construct for transient expression and targeting to the apoplast was then made by ligating the DNA sequence of the signal peptide of *N. benthamiana PR1a*, cloned using the primers PR1a_sp_F and OLC7v35_PR1a_sp_R, to the N-terminus of the *chp-7v35* sequence, cloned using the primers Chp7+a35F and C7+279R, resulting in pCR8_spC7 and pCR8_spC7ST. The entry clones were then recombined into binary vector pMDC32, containing a C-terminal HA-PreScission-Biotin (HPB) epitope (Qi and Katagiri, 2009), through LR reactions (Invitrogen, CA), resulting in pMDC-C7HPB, pMDC-spC7HPB and pMDC-spC7STHPB.

All the constructs derived from PCR products were verified by sequencing. The sequences of the PCR primers are listed in Table 3.

Cloning of *pat-1* homologs from *Cmm* strain Cm15-2.0 sm

To create a linearized plasmid containing the *chp-7* native promoter, pGemC7pro, pRLSTC7 (Syverson 2011) was amplified by PCR using the primers Pro_c7_R and Term_c7_F. *pat-1*, *phpA* and *phpB* were cloned as PCR fragments using KOD Hot Start DNA polymerase (Novagen, CA) with the primers Olc7pro_pat1_F and Olc7term_pat1_R, Olc7pro_phpA_F and Olc7term_phpA_R, and Olc7pro_phpB_F and Olc7term_phpB_R, respectively, using DNA from Cm15-2.0 sm as the template. pGemC7pro_pat1, pGemC7pro_phpA and pGemC7pro_phpB were subsequently constructed by ligating the pGemC7pro vector with the PCR fragment of *pat-1*, *phpA* and *phpB*, respectively, using the Gibson assembly cloning kit (New England Biolabs, MA). The three plasmids were then digested with EcoRI and the fragments were ligated with pHN216 cut open with EcoRI, resulting in pHNC7pro_pat1, pHNC7pro_phpA, and pHNC7pro_phpB.

chpG was cloned as a PCR fragment using KOD Hot Start DNA polymerase (Novagen, CA) with the primers OLPR1a_chpGn38_F and Olc7_chpG_R using DNA from Cm15-2.0 sm as the template, and ligated with the fragment PCR-amplified from pCR8_spC7 with the primers PR1a_R and pCR8attL2_F, resulting in pCR8_spG. This entry clone was then recombined into pEARLEYGATE100 (Earley *et al.* 2006) using the LR reaction (Invitrogen, CA) to generate pEG100_spG. pEG100_spC was cloned following the similar procedure with the starting primers OLPR1a_chpCt33_F and Olc7_chpC_R.

All the clones from Cm15-2.0 sm were verified by sequencing. The sequences of all the clones are identical to the sequences in the reference genome *Cmm* NCPPB 382. The sequences of the PCR primers are listed in Table 3.

Plant assays

Tobacco plants were grown to ten- or eleven-leaf stages in a Conviron growth chamber at 22°C and 75% relative humidity with a 16-hour photoperiod under 100 μM m^{-2} sec^{-1} fluorescent illumination. For complementation assays the seventh and eighth leaves were infiltrated using needleless syringes with *C. michiganensis* strains suspended

in potassium phosphate buffer (3.3 mM K_2HPO_4, 6.7 mM KH_2PO_4, pH 6.5) to an optical density at 600 nm (OD_{600}) = 0.4. The infiltrated leaves were examined for the HR phenotype at 48 hpi, unless otherwise noted. For transient expression assays, *A. tumefaciens* GV3101 strains were suspended to a final OD_{600} = 0.2 in MES buffer (10 mM MES, 10mM $MgSO_4$, pH 5.6) containing 150 μM acetosyringone and incubated at 27°C for 90 minutes before infiltrating into tobacco leaves. The infiltrated leaves were examined for the HR phenotype at 48 hpi.

Expression and purification of Chp-7 from *E. coli*

E. coli BL21 (DE3) pLysS carrying pJEC7H was grown in 20 ml LB broth supplemented with chloramphenicol and kanamycin. Cultures were grown with shaking at 250 rpm at 37°C until early log phase (OD_{600} = 0.2), at which point isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 500 μM to induce *chp-7* expression. Cultures were then transferred to 25°C with continued shaking at 250 rpm for 3 hours. After incubation, the *E. coli* cells were suspended in lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 0.5% (v/v) NP-40, pH 7.6). Tandem histidine-tagged proteins from the cell lysates were pulled down by cobalt-based TALON metal affinity resin chromatography (Clontech, CA). Bound proteins were eluted with imidazole buffer (50 mM KH_2PO_4, 300 mM KCl, 200 mM imidazole, pH 7.8). The resulting eluate was desalted in exchange to potassium phosphate-KCl buffer (3.3 mM K_2HPO_4, 6.7 mM KH_2PO_4, 50mM KCl, pH 6.5) using Quick Spin Protein Columns (Roche, Germany).

Cya assay

The DNA sequence of *cya* was cloned from pDD62 (Casper-Lindley et al., 2002) as a fragment using the primers Cya_OLC7_F and Cya_OLpGEM_R. The fragment contains a 5' overlapping sequence in-frame to the 3' end of *chp-7* and was ligated with linearized pRLSTC7 (Syverson 2011) using the Gibson assembly cloning kit (New England Biolabs, MA). The *chp-7-cya* fusion construct as a plasmid was then digested with EcoRI restriction enzyme and ligated into EcoRI sites of pHN216, resulting in pHNC7Cya, and transformed into competent cells of CmsC7X20. Inoculum was prepared as described above and infiltrated into *N. benthamiana* leaves. Leaf samples were

collected at 0, 10 and 24 hpi and processed to extract cAMP as described (Casper-Lindley et al., 2002). The cAMP concentration was measured with a cAMP ELISA kit (ENZO Life Sciences, NY) following the manufacturer's protocol. Statistical analysis was performed using the lme4 package (Bates *et al*., 2014) in the R environment.

Immunoblots

Protein samples in 1× SDS loading buffer were boiled for 5 minutes, separated on 10% SDS-polyacrylamide gels and blotted to PVDF membranes according to the manufacturer's instructions (Bio-Rad, CA). Membranes were incubated with 5% (w/v) non-fat milk dissolved in TBST buffer (50 mM Tris-HCl, 150 mM NaCl, and 0.1% (v/v) Tween-20, pH 7.6) before incubating with anti-FLAG (clone M2, lot no. 080M6034, Sigma, MO), anti-HA (clone 3F10, lot no. 14034800, Roche, Germany), or anti-biotin antibody (clone D5A7, lot no. 1, Cell Signaling, MA) corresponding to the epitope in samples. After washing the membrane with TBST buffer, the secondary antibody anti-rat IgG horseradish peroxidase (HRP) conjugate (lot no. 9, Cell Signaling, MA) or anti-mouse IgG HRP conjugate (lot no. 0000038165, Promega, WI) was incubated with the membrane. All the antibodies used were diluted according to their manufacturers' recommendations. The blots were developed using the Pierce ECL Western Blotting Substrate (Pierce Biotechnology, IL)

Tables and Legends

Table 1. Bacterial strains used in this study

Strain name/ Index	Description	Source/ Reference
Clavibacter michiganesis subsp. *sepedonicus*		
ATCC 33113/ CIC250	Type strain, wild type, virulent and HR^+	ATCC; Bentley *et al.* 2008
CmsC7X20/ CIC336	ATCC 33113 *chp-7* (CMS_2989)::*cmx*, knockout of *chp-7* created by homologous recombination with pRLSTC7X2; Cm^r	Syverson 2011
CmsC7X20_216	CmsC7X20 transformed with pHN216; Cm^r, Km^r	This study
CmsC7X20_C7	CmsC7X20 transformed with pHNC7; Cm^r, Km^r	This study
CmsC7X20_C7ST	CmsC7X20 transformed with pHNC7ST; Cm^r, Km^r	This study
CmsC7X20_C7-F/ CIC426	CmsC7X20 transformed with pHNC7-F; Cm^r, Km^r	This study
CmsC7X20_C7ST-F/ CIC427	CmsC7X20 transformed with pHNC7ST-F; Cm^r, Km^r	This study
CmsC7X20_C7Cya/ CIC428	CmsC7X20 transformed with pHNC7Cya; Cm^r, Km^r	This study
CmsC7X20_Pat1/ CIC429	CmsC7X20 transformed with pHNC7pro_pat1; Cm^r, Km^r	This study
CmsC7X20_PhpA/ CIC430	CmsC7X20 transformed with pHNC7pro_phpA; Cm^r, Km^r	This study
CmsC7X20_PhpB/ CIC431	CmsC7X20 transformed with pHNC7pro_phpB; Cm^r, Km^r	This study
Clavibacter michiganesis subsp. *michiganensis*		
Cm15-2.0 sm/ CIC14	Representative wild type isolate, spontaneous streptomycin resistant, virulent and HR^+; Sm^r	A. K. Vidaver
CMM100/ CIC389	Cured derivative of wild type *Cmm* NCPPB 382 strain, lacking both native plasmids pCM1 and pCM2, avirulent	Meletzus and Eichenlaub 1991
CMM101chpGβ/ CIC376	Mutant strain created by homologous recombination knock-out *chpG* (CMM_0059) in CMM101 background, lacking native plasmid pCM2; Cm^r	Stork *et al.* 2008

Agrobacterium tumefaciens		
GV3101	GV3101(pMP90); Rf^r, Gm^r	Koncz and Schnell 1986
GV_C7HPB_Cyto/ CIA9	GV3101 transformed with pMDC-C7HPB; Rf^r, Gm^r, Km^r	This study
GV_C7HPB_Apop/ CIA10	GV3101 transformed with pMDC-spC7HPB; Rf^r, Gm^r, Km^r	This study
GV_C7STHPB_Apop/ CIA11	GV3101 transformed with pMDC-spC7STHPB; Rf^r, Gm^r, Km^r	This study
GV_chpG_Apop/ CIA12	GV3101 transformed with pEG100-spG; Rf^r, Gm^r, Km^r	This study
GV_chpC_Apop/ CIA13	GV3101 transformed with pEG100-spC; Rf^r, Gm^r, Km^r	This study
Escherichia coli		
BL21(DE3)	Strain BL21 (DE3) carrying pLysS; Cm^r	Invitrogen
BL_C7H/ CIE239	BL21(DE3) transformed with pJEC7H; Cm^r, Km^r	This study

ATCC, American Type Culture Collection; Cm, chloramphenicol; Km, kanamycin; Gm, gentamicin; Rf, rifampicin; Sm, streptomycin.

Table 2. Foliar responses of *Nicotiana* species infiltrated with *C. michiganensis* subsp. *sepedonicus* and *C. michiganensis* subsp. *michiganensis*

Plant Species	Foliar responses to infiltration							
	***Cms* WT**	***Cms Δchp-7***	**Chp-7**	***Cmm* WT**	***Cmm ΔchpG ΔpCM2***	***Cmm ΔpCM1 ΔpCM2***	**ChpG**	***ChpC***
N. tabacum	HR	ND	HR	HR	C	HR	HR	ND
N. sylvestris	HR	ND	HR	HR	C	HR	HR	ND
N. tomentosiformis	ND	ND	ND	C	C	C	ND	ND
N. benthamiana	ND	-	-	C	C	C	-	-
N. glutinosa	ND	ND	ND	HR	C	HR	HR	ND
N. rustica	C	C	-	C	C	C	-	-
N. glauca	ND	-	-	ND	-	-	-	-
N. clevelandii	ND	-	ND	HR	ND	HR	HR	ND
N. attenuata	ND	-	-	ND	-	-	-	-

Plant foliar responses were recorded at 3 dpi.
WT, wild type; **C,** chlorosis; **HR,** hypersensitive response; **ND,** none detected; **-,** not tested.
Chlorosis symptoms appeared similar to those shown in Figure 5C.

Table 3. PCR primers used in this study

Name	Sequence (5' to 3')
C7_OlpHN_F	GGCCCTTTCGTCTTCAAGAATTCGA
C7_FLAG_R	CTACTTGTCGTCATCGTCTTTGTAGTCCGTTGCTAGGACGTAATAGGGTTGC
FLAG_OlpHN_ R	CTCGCGGACAGGGCCCAGAATTCCTACTTGTCGTCATCGTCTTTGTA
C7_S232T_F	GGGACCTGTCGTCAGTAGGGATA
C7_S232T_R	CCAGTGTCACCCCTCTGCGTCGCTGTTGA
Cya_OLC7_F	TATTACGTCCTAGCAACGCAGCAATCGCATCAGGCT
Cya_OLpGEM_R	GTGATTAAATATTTCACACTATCGATAACTGTCATAGC
Chp7v35_F	ATGGTCGATCGCATCGCCCGTGC
C7+279R	CGTTGCTAGGACGTAATA
PR1a_sp_F	ATGGGATTTGTTCTCTTTTCACA
OLC7v35_PR1a_sp_R	GGGCGATGCGATCGACGGCACGGCAAGAGTGGGATA
Chp7+a35F	GTCGATCGCATCGCCCGTGC
PR1a_R	GGCACGGCAAGAGTGGGATATT
pCR8attL2_F	AAGGGCGAATTCGACCCAGCTT
OLC7KpnI_PIV2_F	CCTGTGCGAGCGGGTACGTAAGTTTCTGCTTCTACCTTTGAT
OLC7KpnI_PIV2_R	CACTCTCACTGAAGATGAGATGGCTGCACATCAACAAATTTTGGTCAT
Term_c7_F	CTGCGGACCGTGACGAA
Pro_c7_R	TAGGGCGAGAGACGGGC
Olc7pro_pat1_F	GCCCGTCTCTCGCCCTAATGCAGTTCATGTCGCGCA
Olc7term_pat1_R	TCGTCACGGTCCGCAGTCAGGAGGTCGCTAATATGT
Olc7pro_phpA_F	GCCCGTCTCTCGCCCTAATGTCGCACATCAGTAGGA
Olc7term_phpA_R	TCGTCACGGTCCGCAGTCAGGAGGTCGCTAATACA
Olc7pro_phpB_F	GCCCGTCTCTCGCCCTAATGAACACAAGCACAAATTCC
Olc7term_phpB_R	TCGTCACGGTCCGCAGTCATCGTCCTGTTACTAGTGC
Olc7_chpC_R	GCTTCGTCACGGTCCGCAGTCAACGTGTCACGAGATTGTATTCT
OLPR1a_chpCt33_F	TATCCCACTCTTGCCGTGCCACCGAGGGCGCCAGGCAGACA
Olc7_chpG_R	GCTTCGTCACGGTCCGCAGTCAGTTGGCGGGTGCGAGCTTGTA
OLPR1a_chpGn38_F	TATCCCACTCTTGCCGTGCCAACGGACTCAGCAACCCGGA

Figures and Legends

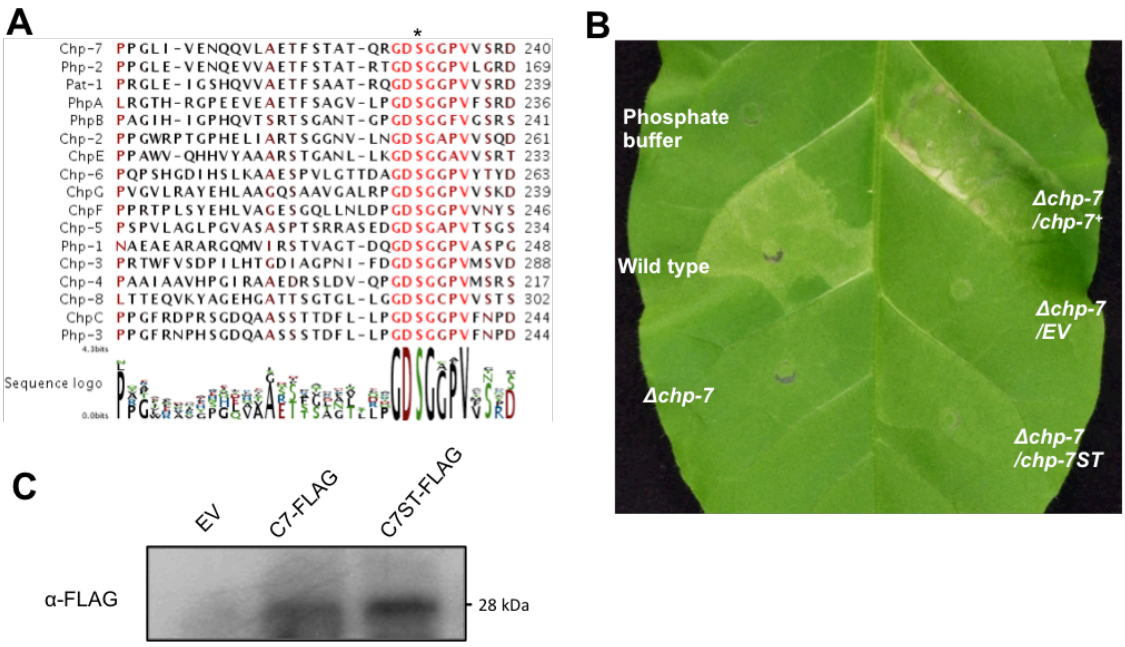


Figure 1. Complementation of the HR-inducing ability of a *C. michiganensis subsp. sepedonicus Δchp-7* mutant in *N. tabacum* with wild type *chp-7* or *chp-7ST*. A, Alignment of the amino acid sequences of *pat-1* homolog genes in *Cmm* and *Cms*. Highly conserved positions are marked in red. Asterisk depicts the Ser residue substituted to Thr in Chp-7ST. **B,** Bacterial suspensions in potassium phosphate buffer were infiltrated into the intercellular spaces of the leaf with needleless syringes. **Wild type** = ATCC 33113; ***Δchp-7*** = CmsC7X20, *chp-7* knockout mutant; ***Δchp-7/EV*** = CmsC7X20_216, CmsC7X20 transformed with empty vector pHN216; ***Δchp-7/chp-7*$^{+}$** = CmsC7X20_C7, CmsC7X20 transformed with pHNC7 carrying *chp-7*; ***Δchp-7/chp-7ST*** = CmsC7X20_C7ST, CmsC7X20 transformed with pHNC7ST carrying *chp-7ST*. The infiltrated leaves were photographed at 48 hpi. Similar results were observed in three independent experiments. **C,** Immunoblot of the supernatants of bacterial cultures in YGM broth with anti-FLAG antibody. Sample inputs EV, C7-FLAG and C7ST-FLAG are the culture supernatants of CmsC7X20_216, CmsC7X20_C7-F and CmsC7X20_C7ST-F, respectively.

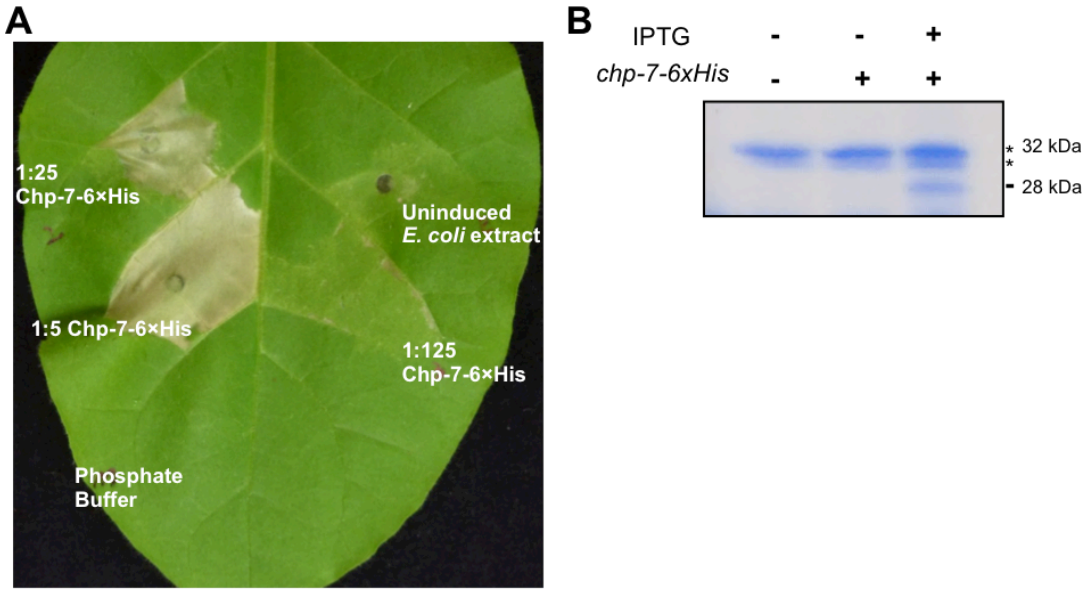


Figure 2. Chp-7-6×His protein purified from *E. coli* is sufficient to trigger the HR. **A,** *N. tabacum* leaves infiltrated with eluent containing purified Chp-7-6×His protein diluted to various extents with potassium phosphate buffer (pH 6.5). A similar undiluted protein extract from an uninduced *E. coli* culture was used as a negative control. The infiltrated leaves were photographed at 48 hpi. The experiment was repeated twice and similar results were observed. **B,** SDS-PAGE of the proteins eluted from metal affinity chromatography resin loaded with cell lysates of *E. coli* strain BL21(DE3) or its derivative transformed with pJEC7H, containing the *chp-7-6xHis* construct, with or without IPTG induction. The protein bands were visualized with Coomassie Brilliant Blue G-250 staining. The predicted mass of Chp-7-6xHis is 28 kDa. Asterisks indicate contaminating proteins from *E. coli.*

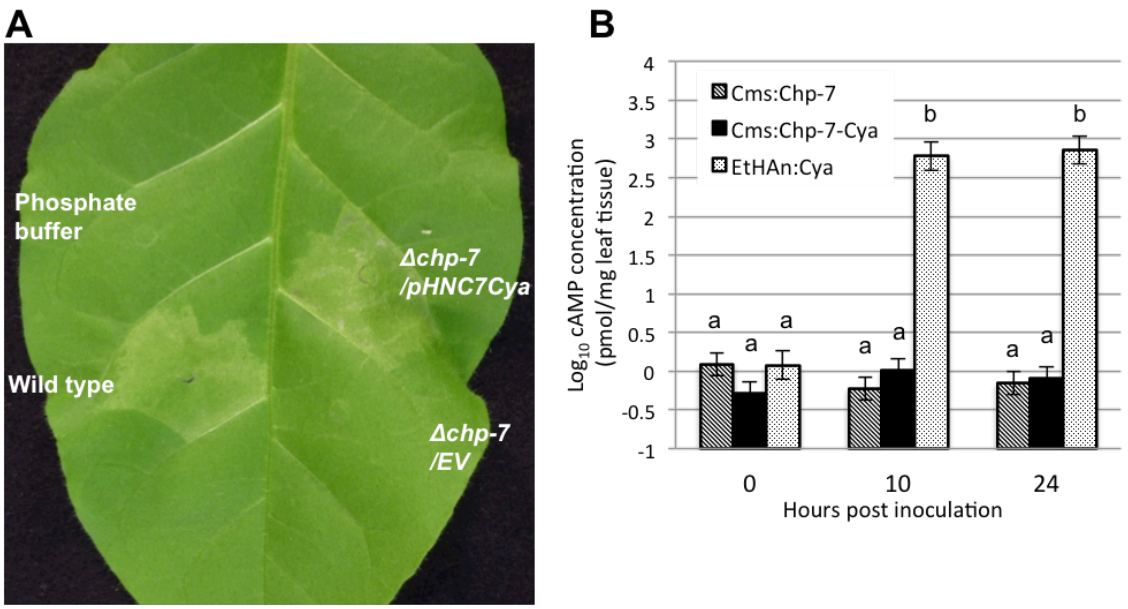


Figure 3. Cya assay of Chp-7 translocation. A, HR in *N. tabacum* elicited by Chp-7-Cya secreted by *Cms*. **Wild type** = ATCC 33113; ***Δchp-7*** = CmsC7X20, *chp-7* knockout mutant; ***Δchp-7/EV*** = CmsC7X20_216, CmsC7X20 transformed with the empty vector pHN216; ***Δchp-7/pHNC7Cya*** = CmsC7X20_C7Cya, CmsC7X20 transformed with pHNC7Cya carrying a *chp-7-cya* fusion. The infiltrated leaves were photographed at 48 hpi. **B,** cAMP measurement in *N. benthamiana* leaf samples infiltrated with *Cms* expressing *chp-7* or *chp-7-cya*, or EtHAn expressing *T3sp-cya* at 0, 10 and 24 hpi. Means and standard errors were estimated from three replicates using a mixed-effect linear model. Different letters indicate significant differences ($p<0.05$).

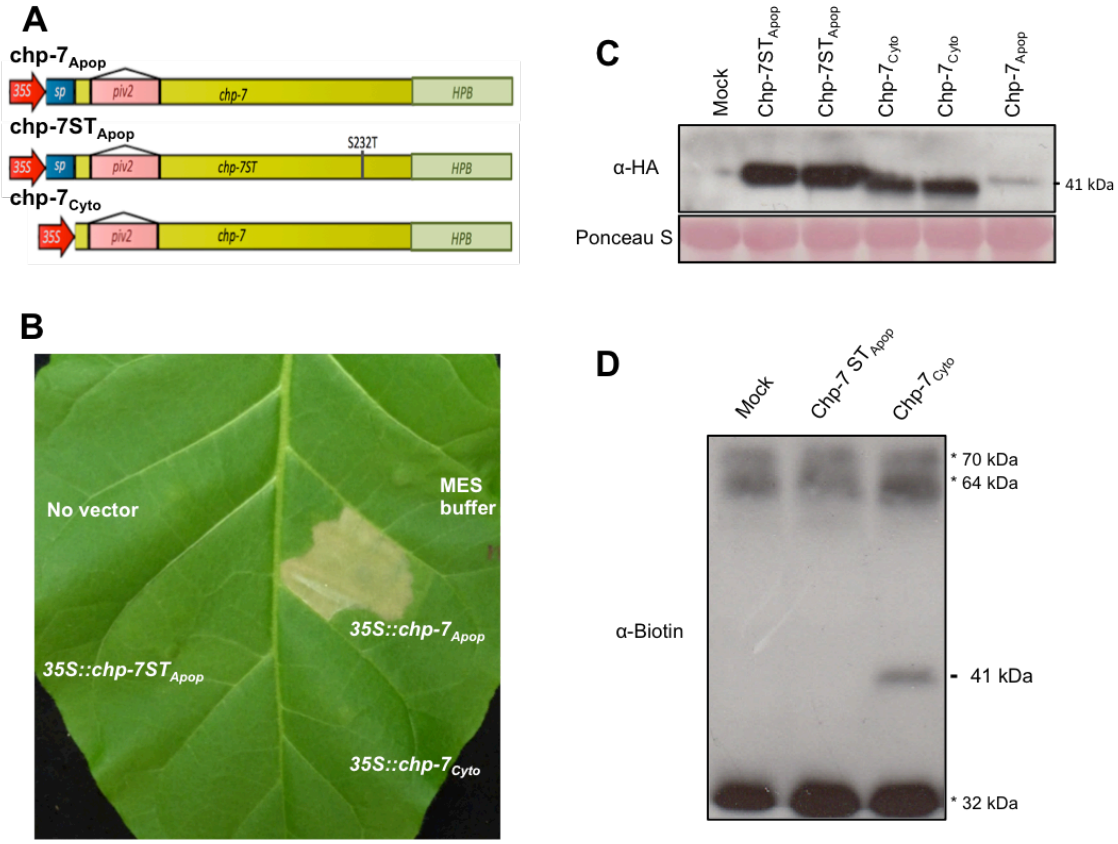


Figure 4. HR elicited by transient expression of *chp-7* targeted to the apoplast of *N. tabacum*. A, Schematics of the *chp-7* transient expression constructs. ***35S***, CaMV *35S* promoter; ***sp*,** signal peptide region of the tobacco *PR1a* gene; ***piv2*,** second intron of the potato *LS-ST1* gene; ***HPB*,** HA-PreScission-Biotin tag. **B,** *N. tabacum* leaf infiltrated with *A. tumefaciens* GV3101 or its derivatives: GV_C7HPB_Apop, carrying ***35S::chp-7$_{Apop}$***; GV_C7STHPB_Apop, carrying ***35S::chp-7ST$_{Apop}$***; or GV_C7HPB_Cyto, carrying ***35S::chp-7$_{Cyto}$***. The infiltrated leaves were photographed at 48 hpi. Similar results were observed in three independent experiments. **C,** Immunoblot with monoclonal anti-HA antibody of total protein extracts from the *N. tabacum* infiltrated leaf areas at 12 hpi. Two independent samples were loaded for *chp-7ST$_{Apop}$* and *chp-7$_{Cyto}$*. The predicted mass of Chp-7$_{Cyto}$ is 41 kDa. Ponceau S staining was used as a loading control. **D,** *in vivo* biotinylation of transiently expressed Chp-7-HPB. Immunoblot of the total protein extracts with monoclonal anti-biotin antibody. Only the sample from leaves expressing

chp-7$_{Cyto}$ shows a band corresponding to biotinylated Chp-7-HPB at approximately 41 kDa. Asterisks indicate endogenous biotinylated proteins in *N. tabacum*.

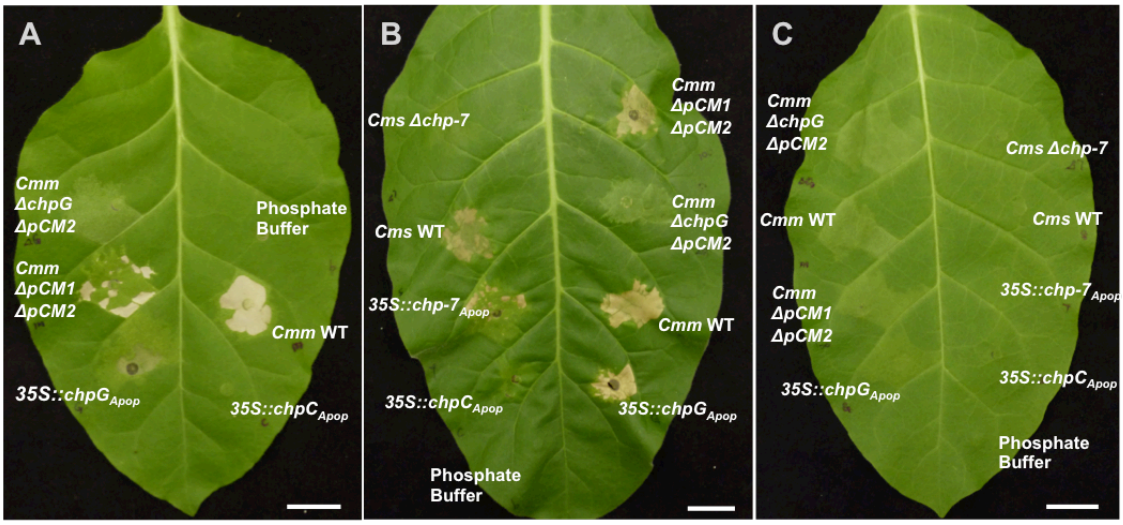


Figure 5. Recognition of Chp-7 and ChpG in *N. tabacum* and one of its progenitor species, *N. sylvestris*. *C. michiganensis* suspensions in potassium phosphate buffer or *A. tumefaciens* suspensions in MES buffer were infiltrated into leaves of **A,** *N. tabacum*; **B,** *N. sylvestris*; and **C,** *N. tomentosiformis.* The bacterial inocula used were ***Cmm* WT =** Cm15-2.0 sm; ***Cmm* ΔchpGΔpCM2** = CMM101chpGβ; ***Cmm* ΔpCM1ΔpCM2 =** CMM100; ***Cms* WT =** ATCC 33113; ***Cms* Δchp-7 =** CmsC7X20; ***35S::chp-7*$_{Apop}$ =** GV_C7HPB_Apop; ***35S::chpG*$_{Apop}$** = GV_chpG_Apop; ***35S::chpC*$_{Apop}$** = GV_chpC_Apop. The infiltrated leaves were photographed at 48 hpi. Similar results were observed in at least six leaves for each experiment. Scale bars = 160 mm.

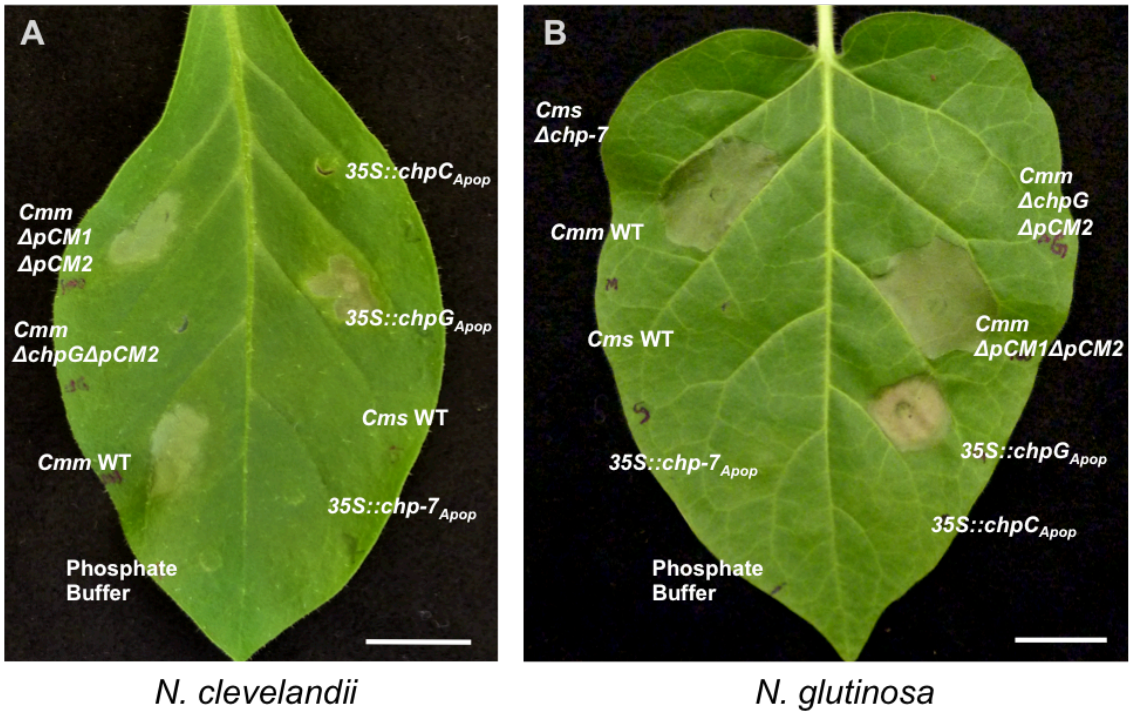


Figure 6. *C. michiganensis* subsp. *michiganensis* elicits *chpG*-dependent HR in *N. clevelandii* and *N. glutinosa*. *C. michiganensis* suspensions in potassium phosphate buffer or *A. tumefaciens* suspensions in MES buffer were infiltrated into leaves of **A,** *N. clevelandii*; and **B,** *N. glutinosa*. The bacterial inocula used were ***Cmm* WT** = Cm15-2.0 sm; ***Cmm ΔchpGΔpCM2*** = CMM101chpGβ; ***Cmm ΔpCM1ΔpCM2*** = CMM100; ***Cms* WT** = ATCC 33113; ***Cms Δchp-7*** = CmsC7X20; ***35S::chp-7*** $_{Apop}$ = GV_C7HPB_Apop; ***35S::chpG*** $_{Apop}$ = GV_chpG_Apop; ***35S::chpC*** $_{Apop}$ = GV_chpC_Apop. The infiltrated leaves of *N. clevelandii* and *N. glutinosa* were photographed at 48 hpi and 72 hpi, respectively. Both experiments were repeated three times with similar results. Scale bars = 160 mm.

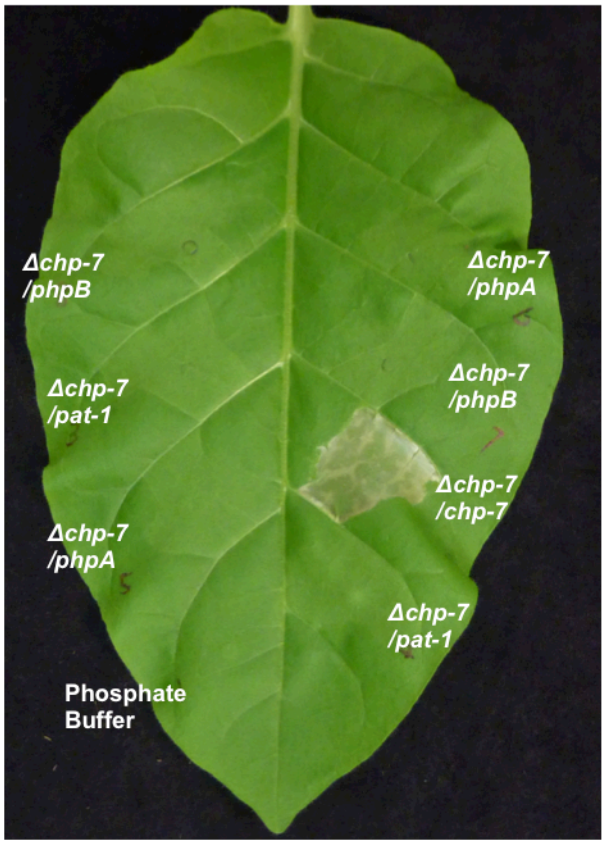


Figure 7. Complementation of the HR-inducing ability of CmsC7X20 in *N. tabacum* with *pat-1*, *phpA* or *phpB* from *C. michiganensis* subsp. *michiganensis*. The *chp-7* knockout strain CmsC7X20 was transformed with the protein coding sequences of *chp-7*, *pat-1*, *phpA* or *phpB* expressed from the *chp-7* promoter. ***Δchp-7/chp-7*** = CmsC7X20_C7; ***Δchp-7/pat-1*** = CmsC7X20_Pat1; ***Δchp-7/phpA*** = CmsC7X20_PhpA; ***Δchp-7/phpB*** = CmsC7X20_PhpB. The infiltrated leaf was photographed at 48 hpi. The experiment was repeated one time and a similar result was observed.

Chapter IV. Characterization of nonhost resistance in *Arabidopsis thaliana* against *Clavibacter michiganensis*

Summary

Nonhost resistance is the most durable form of plant immunity to pathogens. It results from a variety of underlying mechanisms. *Clavibacter michiganensis*, a group of Gram-positive phytopathogens, is non-adapted on plants in the *Brassicaceae*. To understand how plants in the *Brassicaceae* defend themselves against *C. michiganensis*, we screened 32 accessions of *Arabidopsis thaliana* and discovered the responses range from no symptom to chlorosis and necrosis. The allele(s) in *A. thaliana* accession Kas-1, which shows a chlorotic response following *C. michiganensis* inoculation, appears to be dominant over the alleles in Col-0 and Tsu-1, which show no symptom. This variation was controlled by one locus mapped to one 53 kb region on chromosome 4. The ability to induce chlorosis in Kas-1 was conserved among all subspecies of *C. michiganensis* that were tested. Heat-killed and live cells elicited very similar responses. Inoculation with *C. michiganensis* activated MAP kinase cascades in both Kas-1 and Tsu-1. The expression of marker genes for the salicylic acid dependent and *PAD4* dependent signaling pathways was strongly induced by *C. michiganensis* in Kas-1, Col-0 and Ler-0, but not in Tsu-1. However, the titer of *C. michiganensis* persisted in Kas-1 and Ler-0, while it decreased in Col-0 and Tsu-1 to less than 20 percent of the inoculum concentration over a 9-day period following inoculation. These and other results suggest the resistance of *A. thaliana* against *C. michiganensis* may be independent of hormone signaling, and the chlorosis induced by *C. michiganensis* may compromise the resistance. Further investigation will be required to elucidate the molecular basis of the chlorotic response induced by *C. michiganensis* in *A. thaliana*. Knowledge of nonhost resistance responses to *C. michiganensis* can potentially provide new avenues for breeding crops with resistance to this economically important species.

Introduction

Nonhost resistance is a term used to describe the outcomes of some plant-microbe interactions in which all members of one plant species are resistant to all strains of one pathogen species. It has often been observed that one plant pathogen well adapted to one species or a set of closely-related species of plant hosts is not able to infect and cause disease on other distantly related plant species. Therefore, nonhost resistance has been regarded as the most durable form of plant disease resistance (Nurnberger and Lipka, 2005). Although the phenotypes of nonhost resistance are visually similar, the underlying mechanisms may vary and in some cases can be a combination of effects of multiple types of plant defense responses, including preformed resistance, pattern triggered immunity (PTI) and effector triggered immunity (ETI) (Jones and Dangl, 2006).

Non-adapted pathogens can be blocked by certain plant structures and plant secondary metabolites, collectively known as "pre-formed defense mechanisms" (Mysore and Ryu, 2004). For example, oat plants deficient in avenacin A-1 biosynthesis become susceptible to the non-adapted fungal pathogen *Gaeumannomyces graminis* var. *tritici* (Papadopoulou et al., 1999). Nonhost resistance can also be inducible upon pathogen attack. For example, nonhost resistance to some powdery mildew pathogens has been reported to exist at both pre- and post-invasion stages (Lipka et al., 2005). The pre-invasion resistance is conferred by coordinated functions of several *PEN* genes, leading to penetration resistance. In *pen* mutants, the growth of fungal hyphae following successful penetration of plant epidermal cells is blocked by plant defenses mediated by *EDS1*, *PAD4* and *SAG101* (Lipka et al., 2005). This bi-stage model of nonhost resistance has also been observed for some bacterial pathogens. In the *A. thaliana-Pseudomonas* system, the flagellin receptor FLS2 senses the presence of bacteria on the leaf epidermis and mediates stomatal closure. This prevention of entrance through stomata constitutes resistance at the pre-invasion stage (Zeng and He, 2010). Once the stomatal defense has been bypassed, for example by entry through wounds or by artificial infiltration of bacteria into the mesophyll space of leaves, additional inducible plant defense mechanisms, including PTI and ETI, can arrest multiplication of non-adapted pathogens. For example, the *Arabidopsis NHO1* gene contributes to the broad-spectrum suppression of *in planta* growth of *P. syringae* pv. *phaseolicola*, *P. syringae* pv. *tabaci*, *P.*

fluorescens, and a type III secretion system (T3SS) deficient mutant strain of *P. syringae* pv. *tomato* (*Pto*) DC3000 (Lu et al., 2001). *NHO1* contributes to the nonhost resistance against non-adapted *P. syringae* in a flagellin-dependent manner (Li et al., 2005), suggesting FLS2 plays a role in defense beyond stomatal closure. Another example supporting overlap between nonhost resistance and PTI comes from the study of Arabidopsis *CYP710A1* showing this PTI-responsive-gene contributes to restricting the growth of *P. syringae* strains non-adapted to *A. thaliana* by limiting nutrient efflux (Wang et al., 2012). In addition, recognition of T3Es that trigger hypersensitive responses (HR) can cause nonhost resistance in some examples of incompatible interactions. Recognition of HopAS1, a T3E broadly present in *P. syringae,* triggers ETI and contributes to the *in planta* growth inhibition of *Pto* T1 and other *P. syringae* strains non-adapted on *A. thaliana* (Sohn et al., 2012). In contrast, the alleles of *hopAS1* in the *P. syringae* strains that are virulent on *A. thaliana* encode truncated HopAS1, which cannot be recognized and thus avoids triggering ETI (Sohn et al., 2012). Although nonhost resistance dependent on recognition of a certain effector seems easy to break by effector mutations, some effectors, such as HopAS1, may be essential for pathogen virulence and thus remain conserved through co-evolution with hosts. Perhaps the mechanisms underlying nonhost resistance are PTI and/or ETI in many cases, and non-adapted pathogens lack appropriate sets of effectors to suppress PTI or have not evolved to avoid triggering ETI. If so, this might partially explain the observed phenomenon that pathogens of certain plant species cannot multiply and cause diseases on other nonhost species.

Extensive research on the defense signaling network of *A. thaliana* has revealed four signaling sectors: salicylic acid (SA), jasmonic acid (JA), ethylene (ET) and PAD4, which account for up to 80% of the defense output of PTI and ETI (Tsuda et al., 2009). SA is considered the major phytohormone involved in defense against biotrophic pathogens (Glazebrook, 2005). Ca^{2+} signaling contributes to the induction of SA biosynthesis during PTI and ETI, including positive regulation by CBP60g (Wang et al., 2009; Wang et al., 2011), CPKs phosphorylation (Gao et al., 2013), as well as removal of repression by CAMTA3 (Zhang et al., 2014). The initiation of SA signaling is amplified through a feedback loop between the SA and PAD4 sectors (Vlot et al., 2009; Seyfferth

and Tsuda, 2014). SA-triggered plant defense responses are mediated by NPR1, which perceives SA through the activities of NPR3 and NPR4 in an SA concentration-dependent manner (Fu et al., 2012) and functions as the central regulator of downstream transcriptional reprogramming, such as activation of *PR1* (Dong, 2004). JA plays important roles in plant defense against necrotrophic pathogens (Glazebrook, 2005) and is believed to antagonize SA signaling mainly at the transcriptional level (Thaler et al., 2012). The JA receptor COI1 mediates degradation of JAZ proteins, which are transcriptional repressors of JA responsive genes, and activates JA signaling (Thines et al., 2007; Sheard et al., 2010). During plant defense, JA signaling synergistically activates ET signaling through the transcription factors ERF1 and ORA59 (Lorenzo et al., 2003; Pre et al., 2008). However, a recent study using systems biology approaches revealed that JA is also required for SA accumulation and that both JA and PAD4 sectors are inhibited by the ET sector (Kim et al., 2014). It is becoming clearer that the extensive crosstalk between signaling sectors adds to both complexity and robustness of the plant hormone signaling network. Therefore, monitoring the individual activity of the four sectors by sector-specific marker gene expression enables better characterization of the status of the immune signaling network than measuring only the outputs of immune signaling.

The bacterial species *Clavibacter michiganensis* includes six phytopathogenic subspecies, each infecting a narrow range of host plant species, and many unclassified epiphytic or endophytic isolates. The host plants of the phytopathogenic subspecies fall within *Solanaceae*, *Poaceae* and *Fabaceae*. Some examples of nonhost resistance to *C. michiganensis* are associated with HR. *C. michiganensis* subsp. *sepedonicus* (*Cms*) is non-adapted to *Nicotiana tabacum* plants and elicits an HR triggered by one secreted effector, Chp-7 (Nissinen et al., 1997; Nissinen et al., 2009), although it is unclear to what extent Chp-7-triggered HR contributes to growth restriction of *Cms* in *N. tabacum*. Similarly, *C. michiganensis* subsp. *michiganensis* (*Cmm*) elicits a *chpG*-dependent HR in the nonhost plant four-o'clock (*Mirabilis jalapa*) (Stork et al., 2008).

To date, there has been no report of bacterial disease of *Brassicaceae* plants caused by *C. michiganensis*. Therefore, it is likely *C. michiganensis* growth in brassicaceous plants is suppressed by nonhost resistance. *Arabidopsis thaliana*, the model

organism in the *Brassicaceae*, provides an opportunity to investigate nonhost resistance against *C. michiganensis*. In this study, we show that multiple *C. michiganensis* subspecies can induce interveinal chlorosis on several accessions of *A. thaliana*. The chlorotic phenotype appears to be dominant and is mainly controlled by one locus on chromosome 4 of *A. thaliana*. The types of symptoms induced, relative levels of defense gene expression, and rate of pathogen elimination varied among the four *Arabidopsis* accessions in this study.

Results

Clavibacter michiganensis subsp. *michiganensis* induces chlorosis in several *Arabidopsis thaliana* accessions

To determine if *C. michiganensis* can induce any foliar symptoms, such as chlorosis, in *A. thaliana*, 34 accessions were infiltrated with cell suspensions of *Cmm*. The responses of infiltrated leaves were recorded at 6 days post inoculation (dpi). Most of the accessions tested had no visible symptoms, while several accessions responded with mild to severe chlorotic symptoms (Figure 1A; Table 1). Ct-1, Edi-0, Kondara, Ler-0, Rsch-4, Shakdara and Wa-1 developed a mild chlorosis, while An-1, Ob-0, Kas-1 and Kas-2 showed a severe chlorotic to necrotic response. The chlorosis in Kas-1 and Kas-2 became visible as early as 3 dpi, and displayed an interveinal pattern. Thus, *A. thaliana* accessions exhibit various degrees of chlorotic responses to *Cmm*.

The heat stable elicitor of the chlorosis symptom in *Arabidopsis thaliana* accession Kas-1 is conserved among subspecies of *Clavibacter michiganensis*

Approximately 10% of the genes of each *C. michiganensis* subspecies have no homologues in other subspecies (chapter II)(Bentley et al., 2008; Gartemann et al., 2008). Some of these unique genes may have roles in interacting with host plants, contributing to virulence and/or defining the host range. To assess whether the ability to induce the chlorosis symptom in Kas-1 is unique to *Cmm* or conserved in multiple *C. michiganensis* subspecies, inocula prepared from cultures of *Cmm*, *Cms*, *Cmt*, *Cmn* or *Cmi* were infiltrated into the fully expanded leaves of Kas-1. After 6 dpi all subspecies induced

chlorotic symptoms that were indistinguishable from each other (Figure 2). Therefore, the elicitor(s) of the chlorosis could be conserved among *Cmm*, *Cms*, *Cmt*, *Cmn* and *Cmi*.

Expression of many genes related to pathogenicity is inducible upon contact with hosts. To determine whether an elicitor(s) of *C. michiganensis* is produced *in vitro* or only *in vivo*, *A. thaliana* leaves were infiltrated with aliquots of live or heat-treated cells. At 6 dpi, chlorosis was observed in all infiltrated leaves, regardless of heat treatment (Figure 2). No viable cells of *C. michigane*nsis were detected in the heat-treated samples. These results suggest that all five subspecies of *C. michiganensis* tested produce at least one heat-stable elicitor *in vitro*.

The chlorosis induced in *Arabidopsis thaliana* by *Clavibacter michiganensis* is a dominant trait

Among the *A. thaliana* accessions tested, Kas-1 and Kas-2 have the highest sensitivity to *Cmm* inoculation and display severe chlorosis, while Tsu-1 and Col-0 are among the accessions that do not become chlorotic following infiltration with *Cmm*. To analyze the relationship between different alleles of the responsible gene(s), reciprocal crosses between Kas-1 and Tsu-1, and between Kas-1 and Col-0 were conducted to generate F_1 hybrid lines. Leaves of four-week-old F_1 plants were infiltrated with cell suspensions of *Cmm,* and symptoms were assessed at 6 dpi. All F_1 hybrid plants from these crosses showed chlorotic symptoms indistinguishable from Kas-1 (Figure 3). Thus, the Kas-1 allele of the responsible gene(s) appears to be dominant over the Tsu-1 and Col-0 alleles.

The chlorosis induced in *A. thaliana* by *C. michiganensis* is controlled by a single locus on chromosome 4

Given the differences between Kas-1 and Tsu-1 in response to *C. michiganensis* infection, an established population of recombinant inbred lines (RILs) derived from a cross between Kas-1 and Tsu-1 (McKay et al., 2008) was used to dissect the genetic structure controlling the variation. In total 264 RILs were scored for intensity of chlorosis in response to *Cms* using a numeric scale of 0 to 2 (Figure 1B). Each RIL was scored on at least six biological replicates and the average score was used for QTL analysis using

the qtl package in R (Broman et al., 2003). The QTL analysis used 162 markers across the five chromosomes of *A. thaliana* (Figure 4). The significance threshold was calculated using a permutation test, and the 5% significance cut-off was 16.3 in this data set. Using the scanone function with the expectation and maximization (EM) algorithm, one locus of approximately 2.7 Mb flanked by markers Msat4.19 and 4_14385343 on chromosome 4 was detected (Figure 4). This locus was named *Chlorosis by Clavibacter michiganensis 1* (*Ccm1*).

To fine map the *Ccm1* locus, an F_2 population from reciprocal crosses between Kas-1 and Tsu-1 was generated. In total 763 F_2 plants were genotyped with an additional 12 Cleaved Amplified Polymorphic Sequences (CAPS) markers (Konieczny and Ausubel, 1993) present within the *Ccm1* locus (Table 2). The genotyped plants were infiltrated with *Cms* and scored for the chlorosis at 6 dpi. Chlorosis was observed in 553 F_2 plants with various degrees of intensity and 210 F_2 plants showed no visible chlorosis. The segregation of chlorotic to nonchlorotic plants is consistent with the expected 3:1 ratio, with a p-value of 0.108 from a χ^2 test, indicating the chlorotic response is likely controlled only by the *Ccm1* locus. Mapping with the 12 CAPS markers further narrowed down the location of the *Ccm1* locus to an approximately 53 kb region flanked by markers 12.203M_BclI and 12.256M_BsrGI (Figure 5).

Clavibacter michiganensis subsp. *michiganensis* differentially activates plant defense responses in four *Arabidopsis thaliana* accessions

To determine if plant defense signaling is activated by *C. michiganensis* in the four accessions, expression levels of PTI-inducible *CBP60g* and four marker genes (Table 3) reporting the activities of the SA, JA, ET and PAD4 signaling pathways were analyzed using quantitative reverse transcription PCR (qRT-PCR). First, the gene expression changes induced by *Cmm* and *Cms* at 24 hpi in Kas-1 were compared. *CBP60g* and the SA and PAD4 sector reporters were highly induced by both strains, while the JA sector reporter was slightly repressed. The overall trends of gene expression regulation showed good correlation between inoculations with the two subspecies (Figure 6). Then the gene expression changes at 24 hpi in Col-0, Tsu-1, Kas-1 and Ler-0 were examined following inoculation with *Cmm*. Col-0 and Ler-0 showed very similar trends

of strong induction of SA and PAD4 sector activities, suppression of JA sector activity and induction of *CBP60g*, as observed in Kas-1 (Figure 7). Interestingly, up-regulation of ET sector activity was also evident. However, these defense gene inductions were not observed in Tsu-1 at 24 hpi (Figure 7). The two accessions that show chlorosis both have defense genes activated by *Cmm* and *Cms*, and for the two that do not show chlorosis, Col-0 shows similar defense gene induction while Tsu-1 does not. Therefore, there is no obvious correlation between chlorosis symptom and defense gene activation.

MAP kinase (MAPK) activation is an early plant defense response in PTI and ETI. In *A. thaliana*, MPK3, MPK6, MPK4, and MPK11 are the main MAPKs activated in defense responses (Bethke et al., 2012; Meng and Zhang, 2013). MPK11 is the same size as MPK4, and they are not easily distinguished. To test if *Cmm* activates MAP kinases, we conducted immunoblotting with an antibody specifically detecting phosphorylated MAP kinases. MPK3, MPK6 and MPK4/11 were activated in both Kas-1 and Tsu-1 over a 16-hour period post inoculation (Figure 8). The *Cmm*-induced phosphorylation of MAPKs could be detectable by 3-hour post inoculation (hpi) in both Kas-1 and Tsu-1. However, the phosphorylation of MAPKs was sustained in Kas-1, while was reduced to basal levels in Tsu-1 by 16 hpi.

Variation in persistence of *Clavibacter michiganensis* subsp. *michiganensis* among *Arabidopsis* accessions

As *A. thaliana* is considered a nonhost to *C. michiganensis*, the chlorosis symptom observed on some accessions raises the question of if the nonhost resistance is broken, which may permit *C. michiganensis* to persist or to even grow in these accessions. To test this, Kas-1 and Ler-0, which show chlorosis in response to *C. michiganensis*, and Col-0 and Tsu-1, which show no response, were chosen for evaluating *in planta* behavior of *C. michiganensis*. To quantify *C. michiganensis* in leaf tissue, a spontaneous streptomycin resistant *Cmm* strain, Cm15-2.0 sm, was used for isolating bacteria from ground leaves on YGM agar supplemented with streptomycin. Fully expanded leaves of the four *A. thaliana* accessions were infiltrated with cell suspensions of about 5×10^6 cfu/ml *Cmm*. After 9 days, the titer of *Cmm* dropped by almost 10-fold in Tsu-1 and Col-0. However, *Cmm* remained at approximately the same

titer as the initial infiltration dosage in the leaves of Kas-1 and Ler-0 (Figure 9). Therefore, Tsu-1 and Col-0 can reduce the titers of *Cmm* in leaves at a faster rate than Kas-1 and Ler-0. Chlorosis correlates with persistence of *Cmm*, but neither of chlorosis nor persistence correlates with induction of defense gene expression.

Discussion

Nonhost resistance is the most durable form of disease resistance, whose underlying mechanism can vary from preformed defenses to inducible defenses, including PTI and ETI. Understanding how nonhost plants achieve resistance against certain pathogens can benefit breeding efforts to improve crop plant resistance to pathogens. The knowledge gained from studies of nonhost resistance and plant defense responses using model systems has enabled successful engineering of the immune system in crop plants through transferring immune receptors between different plant families (Bouwmeester et al., 2014; Lu et al., 2014; Tripathi et al., 2014; Schoonbeek et al., 2015). Similarly, this work contributes to the understanding of how the nonhost plant *A. thaliana* achieves resistance against the Gram-positive phytopathogen *C. michiganensis*, which is currently causing epidemic diseases on several economically important crops worldwide.

Chlorosis is a typical foliar symptom developed on plants infected with *C. michiganensis* (Larson, 1944). The interveinal chlorosis observed on *A. thaliana* Kas-1 in this study closely resembled the foliar symptom observed on the host plants. It is possible that the chlorosis in *A. thaliana* was caused by virulence activities similar to those used by *C. michiganensis* to infect host plants, though they are not sufficient to overcome the nonhost resistance. Using the model organism *A. thaliana* to study *C. michiganensis* could potentially elucidate the molecular mechanism of the occurrence of this symptom. In fact, *A. thaliana* has been used to study the virulence activities of other Gram-positive phytopathogens. For example, the mode of action of the phytotoxin thaxtomin produced by multiple species of *Streptomyces* was demonstrated using *A. thaliana* (Scheible et al., 2003). In this study, we used a recombinant inbred (RI) population between Kas-1 and Tsu-1 and discovered a single locus controlling the variation in the chlorotic response to *C. michiganensis*. As the two parental lines also differ in defense gene induction by *C.*

michiganensis and rate of reducing *Cmm* population in leaves, this RI population may be useful to map the loci associated with these variations.

Although the identity of the chlorosis elicitor is still unknown at this point, we demonstrated that this elicitor is produced in the cultures of *C. michiganensis* and is heat stable. Some *C. michiganensis* elicitors with these properties are known, such as the glycopeptides from several subspecies *C. michiganensis* that can induce wilting in alfalfa, tomato and potato (Strobel, 1977), as well as the proteinaceous HR elicitor from *Cms* (Nissinen et al., 1997; Nissinen et al., 2009). To elucidate the identity of the elicitor, one approach would be to obtain a mutant strain of *C. michiganensis* that is deficient in producing the responsible elicitor, and test for loss of the ability to induce chlorosis on Kas-1.

As the Kas-1 allele of *Ccm1* is dominant over the Tsu-1 and Col-0 alleles, it is possible that either Kas-1 encodes a target or receptor of the chlorosis elicitor, or the resistance allele of the responsible gene is haplo-insufficient. Certain plant receptors can be exploited by pathogens to promote virulence. Coronatine, a JA-Ile mimicking phytotoxin produced by multiple *P. syringae* strains, requires the JA receptor COI1 to be functional and suppresses SA-dependent defense responses through activation of JA signaling (Xie et al., 1998; Katsir et al., 2008; Zheng et al., 2012). Arabidopsis LOV1, an NB-LRR type R protein that is a susceptibility target of the necrotrophic fungal pathogen *Cochliobolus 88ictoria*, is exploited by the pathogen through victorin-triggered LOV1-mediated resistance-like cell death, leading to disease susceptibility (Lorang et al., 2012). Within the 53 kb *Ccm1* locus, there are 15 genes annotated in the Col-0 reference genome. Among them, there are several possible candidates, such as five genes encoding hypothetical proteins all containing a plant specific domain DUF239 and one R gene encoding a TIR-NB-LRR type R protein. Given that higher than expected recombination frequency and unspecifically mapped sequencing reads of Kas-1 were observed within this region (data not shown), it is possible that Kas-1 or Tsu-1 may encode additional genes not in the Col-0 reference genome. To elucidate the nature of *Ccm1*, it would be necessary to obtain the sequences of the *Ccm1* locus in both Kas-1 and Tsu-1, and compare the gene contents and single nucleotide polymorphisms within the locus from both accessions. The hypothesis that Kas-1 encodes a target or receptor of the chlorosis

elicitor, could be tested by transforming Tsu-1 plants with cloned genes from Kas-1 and examining whether one of the clones alters the response of Tsu-1 to *C. michiganensis*. If so, that cloned gene is likely the responsible gene within *Ccm1*.

Cmm inoculation activates defense signaling in Col-0, Kas-1 and Ler-0, but not in Tsu-1. Intriguingly, *Cmm* titers remain similar to the initial inoculum in the leaves of Kas-1 and Ler-0 even though the plant defense responses were activated, while in both Col-0 and Tsu-1 the population of *Cmm* declined by up to 80% over the 9-day period. It is possible that the mechanism of eliminating *Cmm* is independent of hormone signaling mediated by the four signaling sectors, as Tsu-1 can eliminate *Cmm* without activating plant defense hormone signaling pathways. In this study, we also detected phosphorylation of MAP kinases in both Kas-1 and Tsu-1 over a 16-hour period following inoculation, and the activation was sustained in Kas-1 while transient in Tsu-1 (Figure 8). Indeed, MAPK signaling provides an alternative means for regulating defense responses independent of plant hormone signaling (Tsuda et al., 2013). The roles of MAPK signaling and other plant defense responses, such as ROS production and cell wall reinforcement, in the context of *C. michiganensis* and *A. thaliana* interaction requires further investigation.

The occurrence of chlorosis seems to correlate with persistence of *Cmm* in *A. thaliana* leaves, raising the question of whether chlorosis is the cause or consequence of the persisting *Cmm* population. Given that heat-killed *Cmm* and live *Cmm* triggered similar chlorosis symptoms in Kas-1, it is unlikely that the chlorosis is the consequence of persistent presence of *Cmm* in the inoculated leaves. However, whether the occurrence of chlorosis could directly boost the survival of *Cmm* in the leaves of Kas-1 and Ler-0 is unclear. These questions could be addressed by future research on this topic.

Materials and Methods

Arabidopsis accessions and generation of mapping populations

All *Arabidopsis thaliana* accessions and the Kas-1×Tsu-1 RILs (CS97026) were ordered from the Arabidopsis Biological Resource Center (ABRC) at the Ohio State University. All plants were grown in 12 cm (length) × 12 cm (width) × 5 cm (depth) pots in Berger BM2 germinating mix placed in a Conviron growth chamber at 22°C and 75%

relative humidity with a photoperiod of 12 hours light and 12 hours dark under 100 μM m^{-2} sec^{-1} fluorescent illumination. Information about all the *A. thaliana* accessions used in this study is summarized in Table 1.

Bacterial strains and culture conditions

The *C. michiganensis* strains used in this study are described in Table 4. All strains were cultured on YGM agar (De Boer and Copeman, 1980) at 25.5°C for 5 days before use. The *Cmm* strain is a spontaneous mutant resistant to streptomycin and was therefore cultured on YGM agar supplemented with 50 μg/ml streptomycin.

Plant chlorosis assay and *in planta C. michiganensis* survival assay

To prepare the inocula for *C. michiganensis* inoculation of *A. thaliana*, cultures of *C. michiganensis* subspecies were suspended in inoculation buffer (1 mM $MgSO_4$, 0.5 mM KCl, 0.174 mM KH_2PO_4 and 0.026 mM K_2HPO_4, pH 6.0) to OD_{600}=0.1 which contains approximately 1×10^8 cells. The inocula were infiltrated into the leaves of 30-day-old plants from the underside using needleless syringes. Infiltrated plants were maintained under the same conditions as used for plant growth, and symptoms were assessed over a 6-day period.

The *in planta C. michiganensis* survival assay was performed similarly, but using *Cmm* at OD_{600}=0.005. The inoculated leaves were collected at 0 and 9 dpi with a cork borer with a diameter of 6 mm. To quantify the population of *Cmm* in the samples, the collected leaves were soaked in the inoculation buffer and homogenized with a metal bead under vigorous shaking. The homogenized mixture was diluted and spread on YGM agar supplemented with 50 μg/ml streptomycin. The colony forming units (CFUs) were counted after the plates had been incubated at 25.5 °C for 3 days.

Mapping of the *Ccm1* locus

The genetic structure of the chlorosis response to *C. michiganensis* was first analyzed with an established RI population (CS97026) originated from the parental lines Kas-1 (CS903) and Tsu-1 (CS1640). Each leaf sample was scored on a scale of 0 to 2 for intensity of chlorosis. A score of 0 corresponds to no chlorosis, as observed on the Tsu-1

accession; 1 corresponds to mild chlorosis; and 2 corresponds to severe chlorosis, as observed on the Kas-1 accession (Figure 1B). Each RI line was scored on two leaves per plant for at least three plants, and the average score for the RI line was used as the phenotype score to fit into QTL analysis using the scanone function in the R package qtl (Broman et al., 2003). The genotyping data used for this analysis were obtained from the Arabidopsis Information Resource (TAIR) website (ftp://ftp.arabidopsis.org/home/tair/Maps//Kas_Tsu_RIdata.xls).

The *Ccm1* locus was then mapped using an F_2 population generated from reciprocal crosses between Kas-1 (CS903) and Tsu-1 (CS1640). The F_2 plants were phenotyped based on the occurrence of chlorosis at 6 dpi following *C. michiganensis* inoculation and were genotyped using the 12 CAPS markers (Table 2) within the region bordered by the markers Msat4.19 and 4_14385343.

Gene expression analysis

Fully expanded leaves from 30-day-old *A. thaliana* plants were infiltrated with inoculation buffer or *Cmm* suspended in the inoculation buffer at the concentration of 2×10^7 cfu/ml. Leaf samples were collected at 6 or 24 hpi and homogenized while frozen in liquid nitrogen. Messenger RNA was extracted from the homogenized samples using the TRIzol Reagent (Life Technologies, Carlsbad, CA, USA) following the manufacturer's recommended protocol. The resulting mRNA pellets were dissolved in diethylpyrocarbonate (DEPC) treated water and adjusted to 15 ng/µl. The abundance of gene transcripts of interest was quantified through quantitative reverse-transcription PCR (qRT-PCR) using the SuperScript III Platinum SYBR Green One-Step qRT-PCR Kit (Life Technology, Carlsbad, CA, USA) in a Roche LightCycler 480 System (Roche Applied Science, Germany). Primer sequences for *JAZ10* (At5g13220), *ACS6* (At4g11280), *PR1* (At2g14610), *CBP60g* (At5g26920), At5g46960 and *ACTIN2* (At3g18780) are listed in Table 3. Calculation of gene expression was done using the second derivative maximum method for Cp calling in the LightCycle 480 software (version 1.5.0 SP3) and normalized to expression of *ACTIN2*.

Tables and Legends

Table 1. Responses of 34 *Arabidopsis thaliana* accessions to *Clavibacter michiganensis* subsp. *michiganensis*

Accession Name	Stock Number	Country	Location	Response type
An-1	CS944	Belgium	Antwerpen	Chlorosis and necrosis
Bay-0	CS57923	Germany	Bayreuth	No response
Bur-0	CS6643	Ireland	Burren	No response
Can-0	CS28130	Spain	Canary Islands	No response
Col-0	CS6673	USA	Columbia	No response
Cvi-0	CS902	Cape Verdi	NA	No response
Ct-1	CS6674	Italy	Catania	Mild chlorosis
Edi-0	CS6688	United Kingdom	Edinburgh	Mild chlorosis
Ei-2	CS1124	Germany	Eifel	No response
Fei-0	CS9995	Portugal	St. Maria d. Feiria	No response
Hi-0	CS6736	Netherlands	Hilversum	No response
Kas-1	CS903	India	Kashmir	Chlorosis and necrosis
Kas-2	CS1264	India	Kashmir	Chlorosis and necrosis
Kn-0	CS6762	Lithuania	Kaunas	No response
Kondara	CS6175	Tadjikistan	Khurmatov	Mild chlorosis
Ler-0	CS20	Germany	NA	Mild chlorosis
Mt-0	CS1380	Libya	Martuba/Cyrenaika	No response
Nd-1	CS1636	Germany	Niederzenz	No response
No-0	CS6805	Germany	Halle	No response
Ob-0	CS6816	Germany	Oberursel/Hasen	Chlorosis and necrosis
Oy-0	CS6824	Norway	Oystese	No response
Po-0	CS6839	Germany	Poppelsdorf	No response
Rsch-4	CS6850	Russia	Rschew/Starize	Mild chlorosis
Sav-0	CS6856	Czech	Slavice	No response

		Republic		
Sf-2	CS6857	Spain	San Feliu	No response
Shakdara	CS22652	Tadjikistan	Pamiro-Alay	Mild chlorosis
St-0	CS38906	Sweden	Stockholm	No response
Tsu-0	CS1640	Japan	Tsushima	No response
Var2-1	CS28798	Sweden	Varhallarna	No response
Wa-1	CS6885	Poland	Warsaw	Mild chlorosis
Wil-2	CS6889	Russia	Wilna/Litvanian	No response
Ws-0	CS6891	Russia	Wassilewskija	No response
Wu-0	CS6897	Germany	Wurzburg	No response
Zu-0	CS6902	Switzerland	Zurich	No response

NA, not available.

Table 2. Summary of the CAPS markers used for mapping *Ccm1*

CAPS Marker		Primer Sequence (5' to 3')	Pattern after enzyme digestion (bp)	
			Kas-1	Tsu-1
4_11.782M_HindIII	F	ACCACGAGCTTGCGATTTCTT	406	188+218
	R	AGATTGATCACGCCTGAAAGGT		
4_12.093M_AflII	F	CCACAAAATCTGACGGTCCCA	490+256	746
	R	AAAGGCCGTGATAGAGACGA		
4_12.185M_NsiI	F	GGAGAAACGGCTCGCAACTAA	78+695 +333	78+1025
	R	AACCGAATGAGGCATAACCGA		
4_12.203M_BclI	F	TGTAACAGTTTTGATTGGTCTGGT	346+299	645
	R	GCTCGTTTATATGCCTCTTGGA		
4_12.215M_AlwNI	F	AGCATCCTCCAGGACCTCCAA	307+516	823
	R	GTTAGTCTTGCAACCATGGCG		
4_12.225M_HindIII	F	TAGTCATTGAGGAGGCCCACA	139+278	417
	R	CTTGGCCATTTGGGGAAACCT		
4_12.256M_BsrGI	F	TGAGAGGGATTTGCCCCCATT	331	198+133
	R	AATGGGCTTTGGCCGATTTCA		
4_12.270M_EcoRI	F	AAACACCAGACGATCAGTGCC	582	211+371
	R	GGGAGGCCAATCACAACACTC		
4_12.496M_NsiI	F	CAAACCTTAGCCCTTTTGACCGT	707	330+376
	R	GCAAAAGATGCTGCATTACGATGA		
4_12.638M_EcoRI	F	AGTGGCACCAAAAGTTGCACA	475+714	1189
	R	TGCTCGTAGTGCAGGGTGTAA		
4_13.036M_ClaI	F	GCCACTGCTCCTCTGCATATT	1276	718+558
	R	ATGTTGTGTGTTGTGGACGAG		
4_13.986M_XhoI	F	TGGCGAACAATGAGGGAGAGA	245+493	738
	R	CGAATGCAGCTGGTGTCAACT		

F, forward primer; **R,** reverse primer.

Table 3. Summary of the primers used for qRT-PCR analysis of gene expression

Gene	Primer Sequence (5' to 3')		Reporter for
CBP60g	Forward	CGGGCGTAACACTTCTCTTC	PTI
	Reverse	AGCTTCGGCCTTTAATTGGT	
JAZ10	Forward	ATCCCGATTTCTCCGGTCCA	JA
	Reverse	ACTTTCTCCTTGCGATGGGAAGA	
PR1	Forward	CGGAGCTACGCAGAACAACT	SA
	Reverse	CTCGCTAACCCACATGTTCA	
At5g46960	Forward	GAAAGACCCGCAATTGTCAT	PAD4
	Reverse	CGTCGATGCTAGGACCAAAC	
ACS6	Forward	TCACCAAGTGAAGCTCAACG	ET
	Reverse	CATAGTTGTTGCAGCCATCG	
ACTIN2	Forward	AGTGTCTGGATCGGTGGTTC	Internal Reference
	Reverse	CCCCAGCTTTTTAAGCCTTT	

PTI, pattern-triggered immunity; **JA,** jasmonic acid signaling; **SA,** salicylic acid signaling; **PAD4,** *PAD4*-dependent signaling; and **ET,** ethylene signaling.

Table 4. *Clavibacter michiganensis* strains used in this study

Strain Name	Strain Index	Taxonomy	Source/ Reference
Cm15-2.0 sm	CIC014	*C.m.* subsp. *michiganensis*	A. K. Vidaver
ATCC 33113	CIC250	*C.m.* subsp. *sepedonicus*	ATCC
82055.9	CIC021	*C.m.* subsp. *tessellarius*	A. K. Vidaver
CMN1	CIC251	*C.m.* subsp. *nebraskensis*	C. Ishimaru
ATCC 10253		*C.m.* subsp. *insidiosus*	ATCC

ATCC, American Type Culture Collection.

Figures and Legends

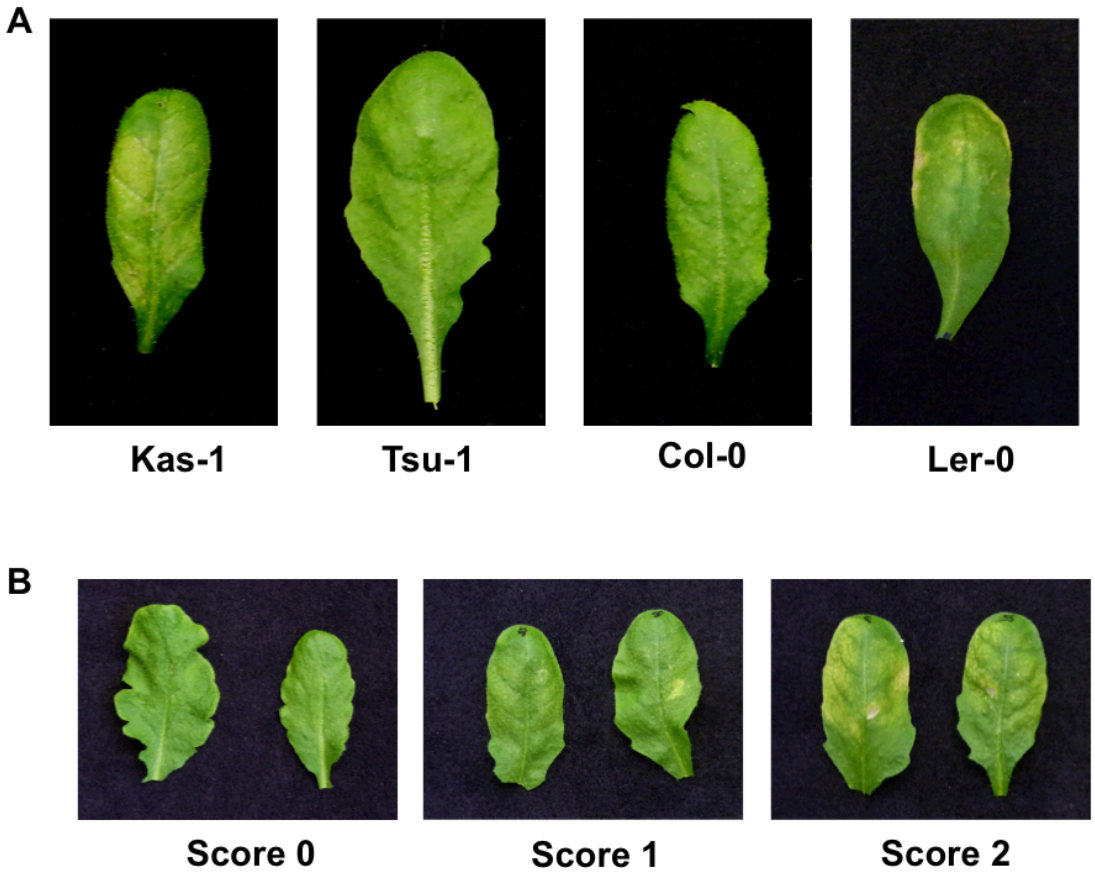


Figure 1. Foliar responses of *Arabidopsis thaliana* accessions to *Clavibacter michiganensis* subsp. *michiganensis* inoculation. A, responses of Kas-1, Tsu-1, Col-0 and Ler-0 to *Cmm* inoculation. **B,** Scoring scheme used for describing the intensity of the responses. **Score 0,** no visible response; **Score 1,** mild chlorosis; and **Score 2,** severe chlorosis and necrosis. All photos were taken at 6 dpi.

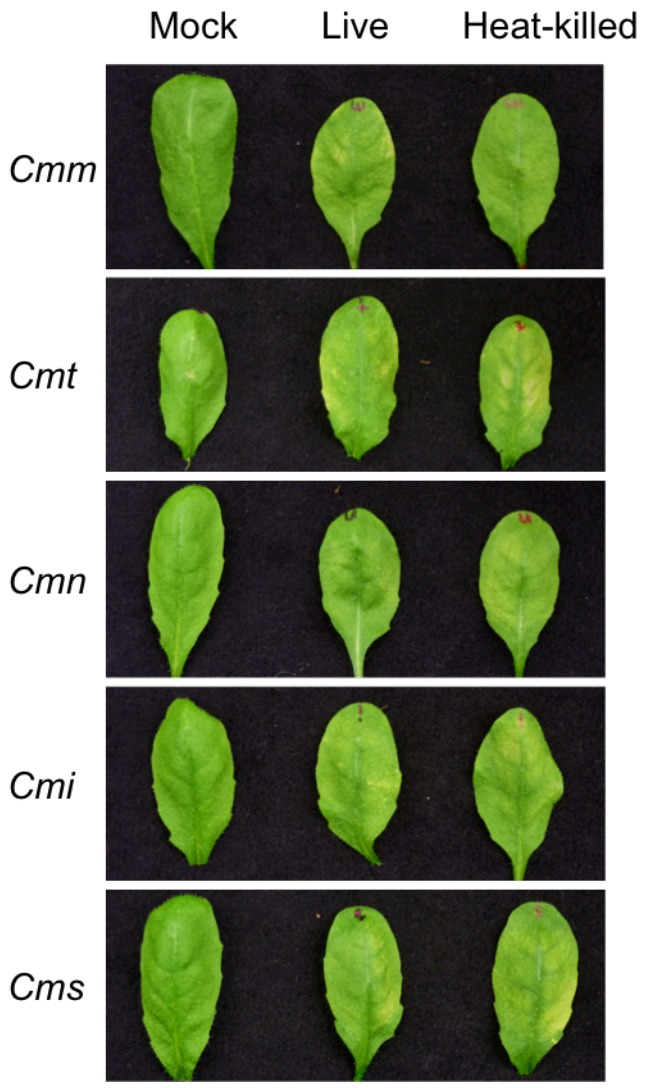


Figure 2. Foliar response of *A. thaliana* Kas-1 to different subspecies of *C. michiganensis*. Leaves of 30-day-old *A. thaliana* Kas-1 plants were infiltrated with buffer or cells suspensions containing approximately 1×10^8 cfu/ml either live or heat-killed ***Cmm*****,** *C. michiganensis* subsp. *michiganensis*; ***Cmt*****,** *C. michiganensis* subsp. *tessellarius*; ***Cmn*****,** *C. michiganensis* subsp. *nebraskensis*; ***Cmi*** *C. michiganensis* subsp. *insidiosus*; or ***Cms*****,** *C. michiganensis* subsp. *sepedonicus*. Photos were taken at 6 dpi.

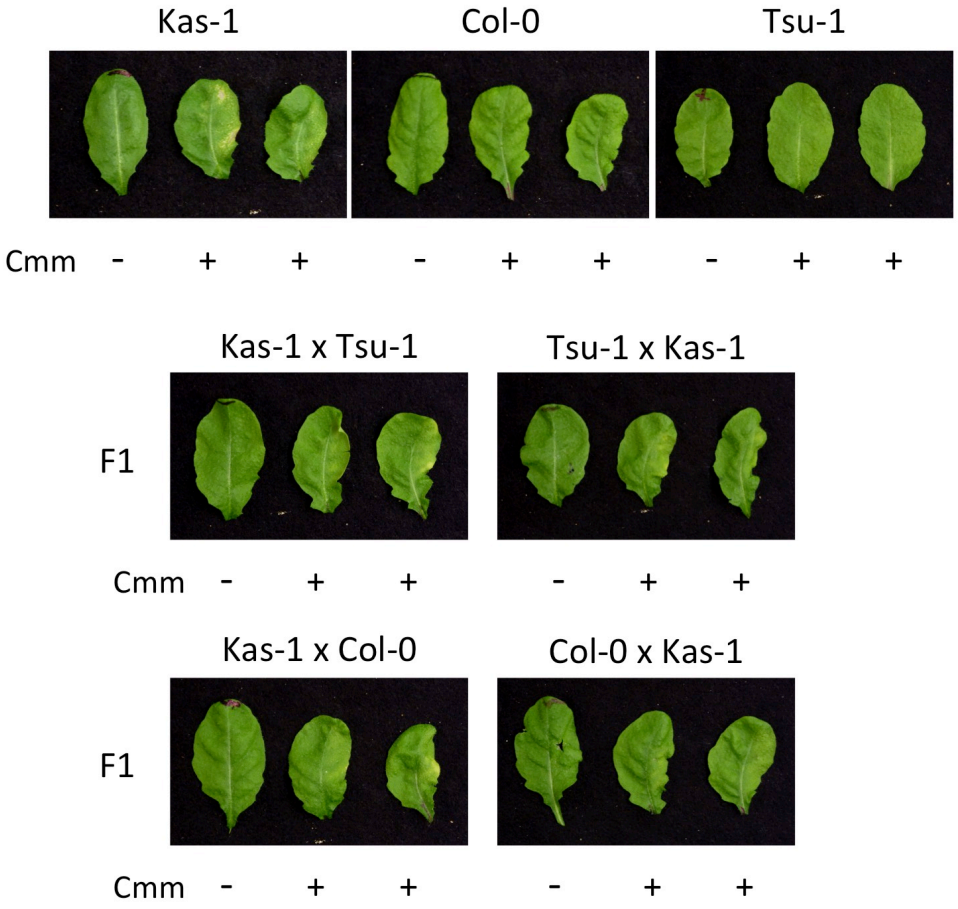


Figure 3. Foliar response of *A. thaliana* F_1 hybrids to *C. michiganensis* subsp. *michiganensis*. Leaves of 30-day-old *A. thaliana* plants were infiltrated with either the inoculation buffer or cell suspensions of *C. michiganensis* subsp. *michiganensis* (**Cmm**) at approximately 1×10^8 cfu/ml. **Top,** the three parental accessions, Kas-1, Col-0 and Tsu-1; **Middle,** F_1 plants from reciprocal crosses between Kas-1 and Tsu-1; **Bottom,** F_1 plants from reciprocal crosses between Kas-1 and Col-0. Photos were taken at 6 dpi.

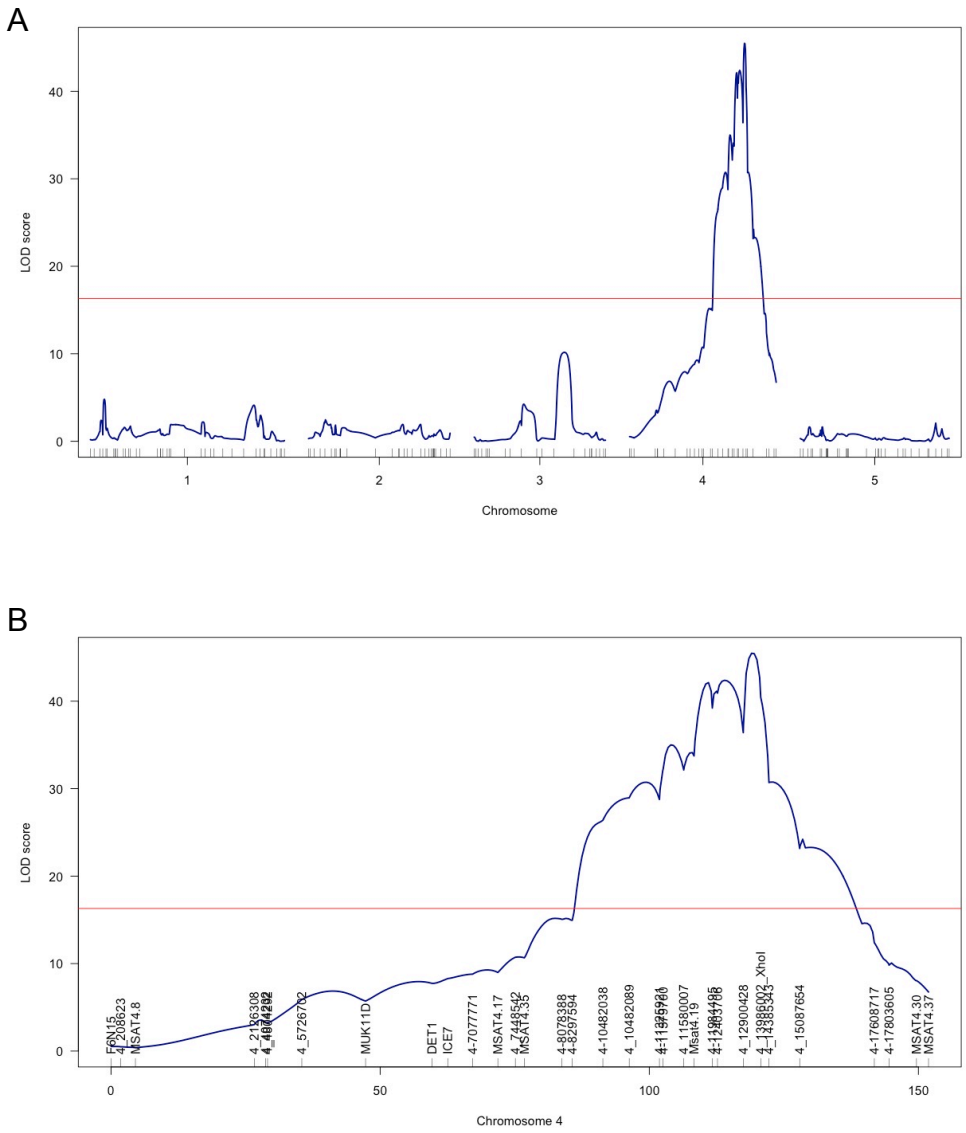


Figure 4. QTL analysis of the chlorotic response in *A. thaliana* Kas-1×Tsu-1 recombinant inbred lines. The Kas-1×Tsu-1 RI lines were scored for chlorosis in response to inoculation with *Clavibacter michiganensis* subsp. *sepedonicus*. The score for each RI line was averaged among at least four replicates. One-QTL scanning was conducted using the EM algorithm in the qtl package in the R environment. The LOD scores (x-axis) along the chromosome positions were plotted against **A,** all five chromosomes, and **B,** chromosome 4. Significance threshold of LOD score at p=0.05 was determined through a permutation test, and are indicated as red lines. Marks on the y-axis depict the positions of the markers used in the QTL analysis.

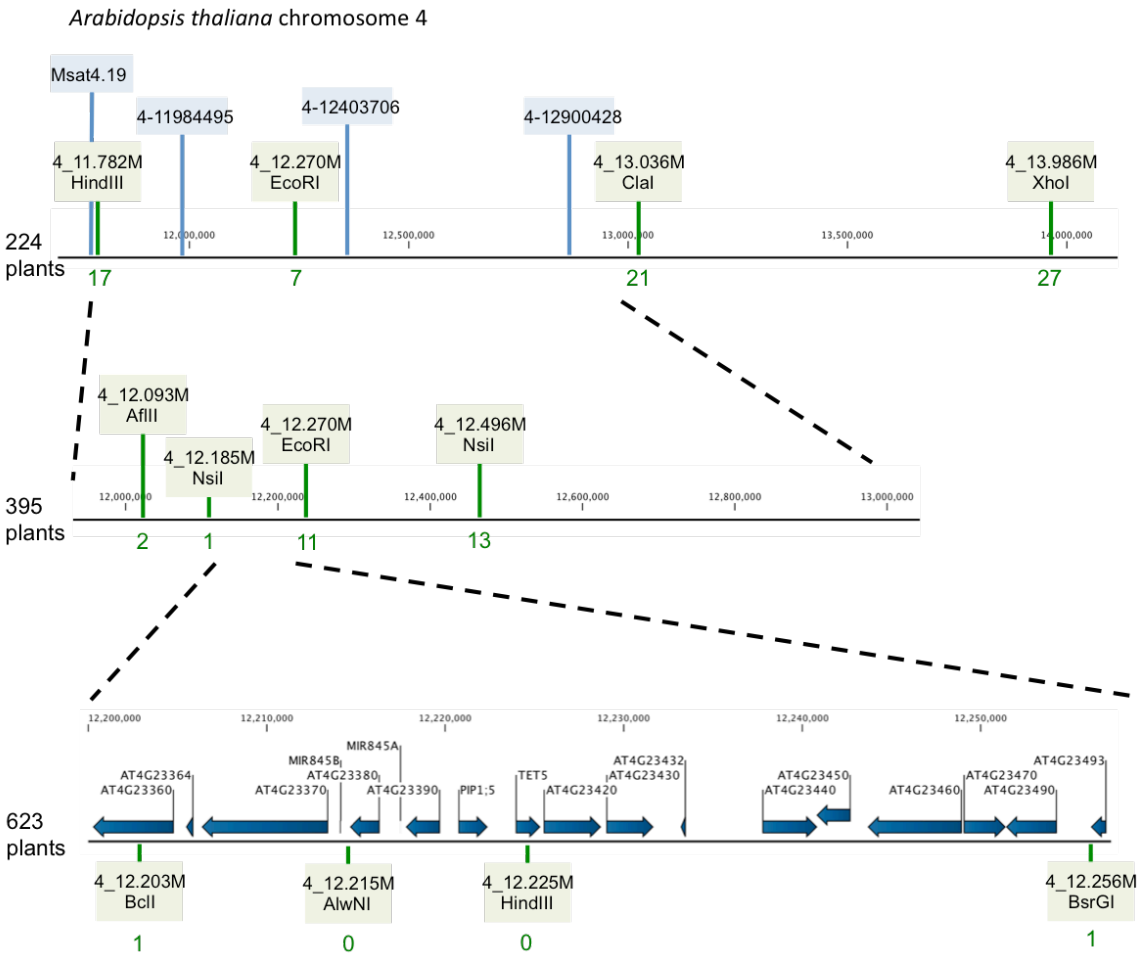


Figure 5. Fine mapping of the *Ccm1* locus in an F_2 population from a cross between Kas-1 and Tsu-1. The F_2 progeny from reciprocal crosses between Kas-1 and Tsu-1 were phenotyped for chlorosis on a scale of 0-2 following infiltration with *Clavibacter michiganensis* subsp. *sepedonicus* and genotyped with a set of CAPS markers (Table 2) within the region bordered by markers Msat4.19 and 4-14385343. The CAPS markers are shown in green, and the markers used during the mapping with RILs are shown in blue. The number under the position of each CAPS marker indicates the number of observed recombination events between phenotype and marker genotype among the number of plants indicated on the left.

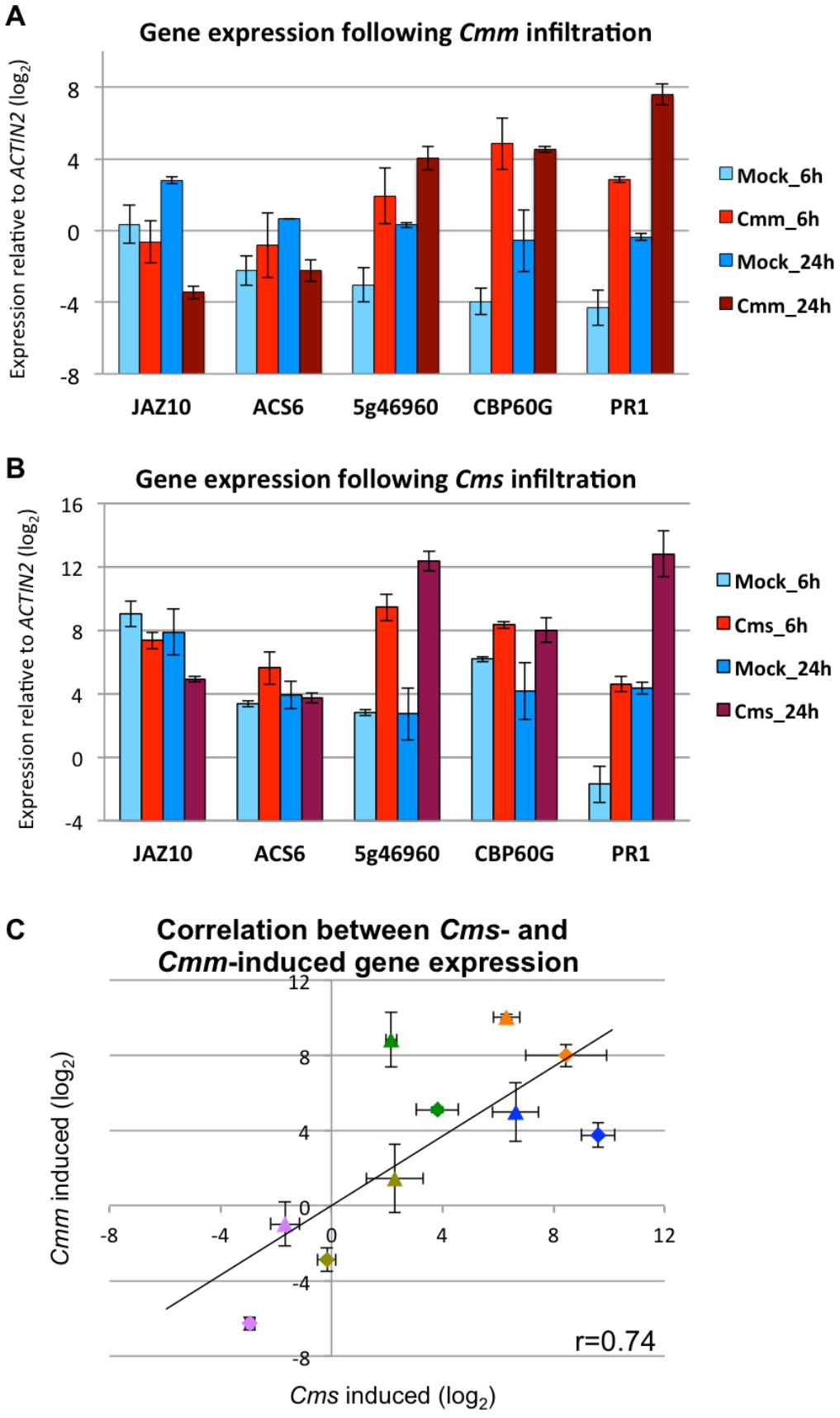

A
Gene expression following Cmm infiltration
Expression relative to ACTIN2 (log2)
8
4
0
-4
-8
JAZ10
ACS6
5g46960
CBP60G
PR1
Mock_6h
Cmm_6h
Mock_24h
Cmm_24h
B
Gene expression following Cms infiltration
Expression relative to ACTIN2 (log2)
16
12
8
4
0
-4
JAZ10
ACS6
5g46960
CBP60G
PR1
Mock_6h
Cms_6h
Mock_24h
Cms_24h
C
Correlation between Cms- and Cmm-induced gene expression
Cmm induced (log2)
Cms induced (log2)
12
8
4
0
-4
-8
-8
-4
0
4
8
12
r=0.74

Figure 6. Correlation between *C. michiganensis subsp. michiganensis* and *C. michiganensis subsp. sepedonicus*-induced gene expression change. Leaves of 30-day-old Kas-1 plants were infiltrated with mock buffer or inoculum containing approximately 2×10^7 cfu/ml of **A,** *Cmm* or **B,** *Cms*. **A** and **B,** relative gene expressions at 6 and 24 hpi were measured with quantitative RT-PCR. **C,** correlation between *Cmm* and *Cms*-induced gene expression at 6 and 24 hpi. The fold inductions (inoculated vs. mock, $\log_2$) of *JAZ10* (shown in purple), *ACS6* (in yellow), *At5g46960* (in blue), *CBP60g* (in green) and *PR1a* (in orange) and by *Cmm* inoculation (x-axis) are plotted against the fold induction of the same genes by *Cms* inoculation (y-axis). Triangle and diamond symbols indicate 6 and 24 hpi, respectively. The displayed r^-value is the coefficient of Pearson correlation. Error bars indicate standard deviation of the mean from two replicates for each data point.

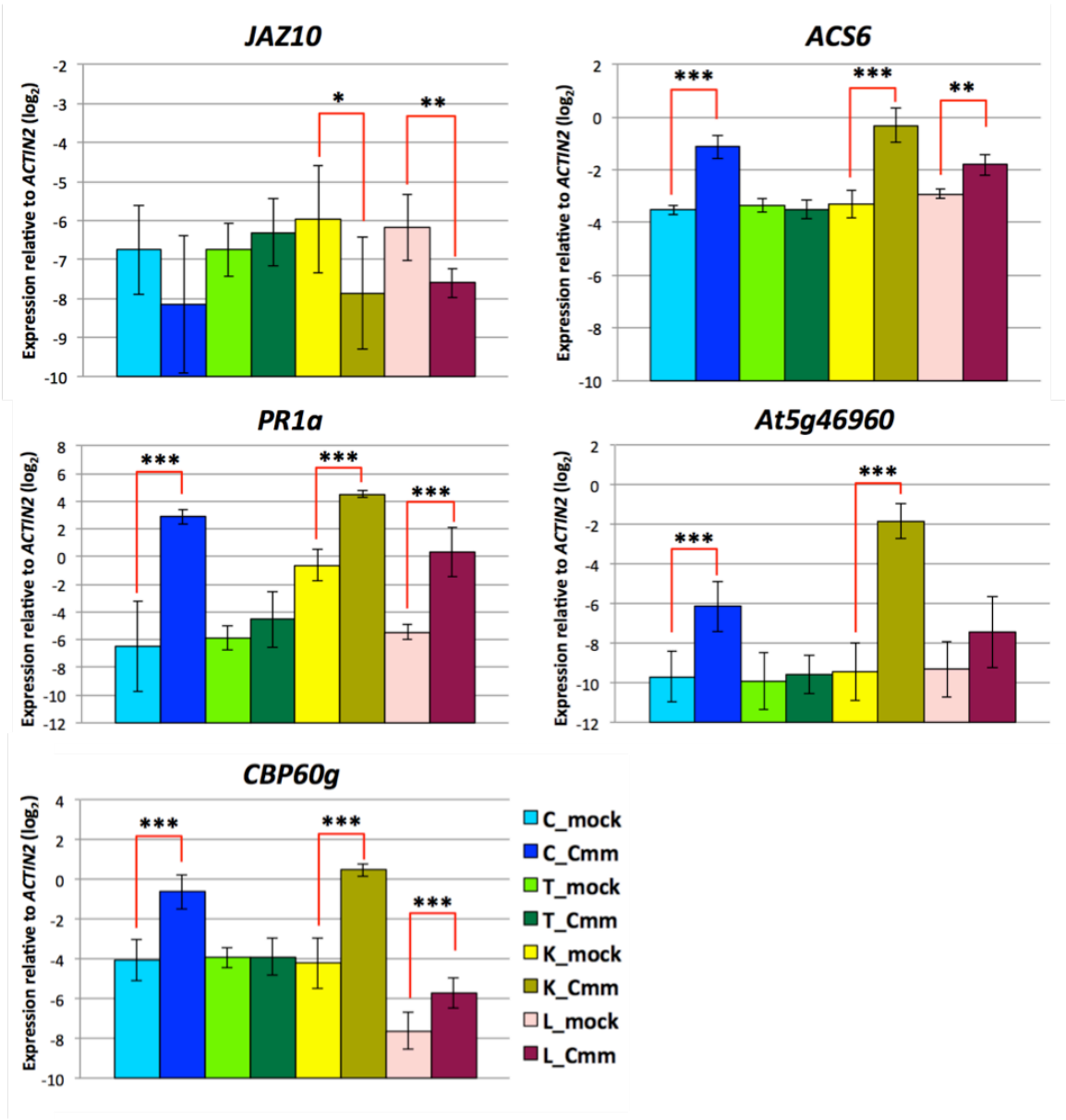


Figure 7. Gene expression change induced by *C. michiganensis* subsp. *michiganensis* inoculation. Leaves of 30-day-old *A. thaliana* accessions **C,** Col-0; **T,** Tsu-1; **K,** Kas-1; **L,** Ler-0, were infiltrated with **mock,** the inoculation buffer; or **Cmm,** cell suspensions of *Cmm* at OD_{600}=0.02, containing approximately 2×10^7 cfu/ml. Leaf samples were collected at 24 hpi. The relative gene expression levels of *JAZ10*, *ACS6*, *CBP60g*, *PR1a* and *At5g46960* were normalized to expression of *ACTIN2*. The error bars represent standard deviations of the mean of n=6 and n=8 biological replicates each with two technical replicates for mock treated samples and for *Cmm* treated samples, respectively. Student's t-test was performed to compare the relative expression level between mock

and *Cmm* treated leaf samples from the same accession. *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$.

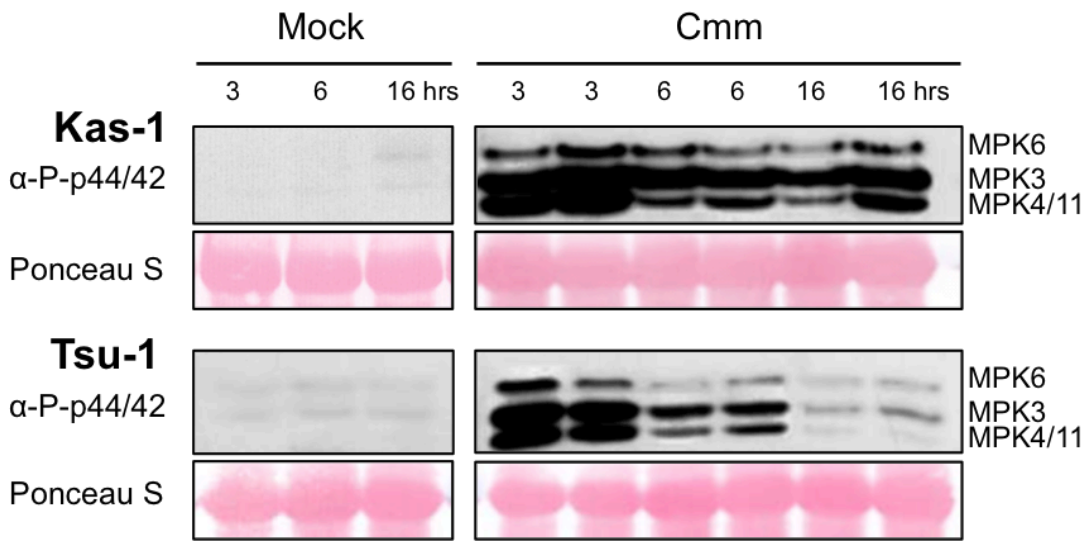


Figure 8. Immunoblot of phosphorylated MAP kinases in Kas-1 and Tsu-1 following inoculation with *C. michiganensis* subsp. *michiganensis*. Leaves of 30-day-old plants were infiltrated with **Mock,** the inoculation buffer; or **Cmm,** cell suspensions of *Cmm* at OD_{600}=0.02, containing approximately 2 x10^7 cfu/ml. Samples were collected at the indicated time points post inoculation. Phosphorylated MPK6/3/4/11 were detected by immunoblotting with anti-Phospho-p44/42 MAPK antibody. MPK4 and MPK11 have similar size and appear as one band in gel. The PVDF membranes were stained with Ponceau S to show equal loading of each lane.

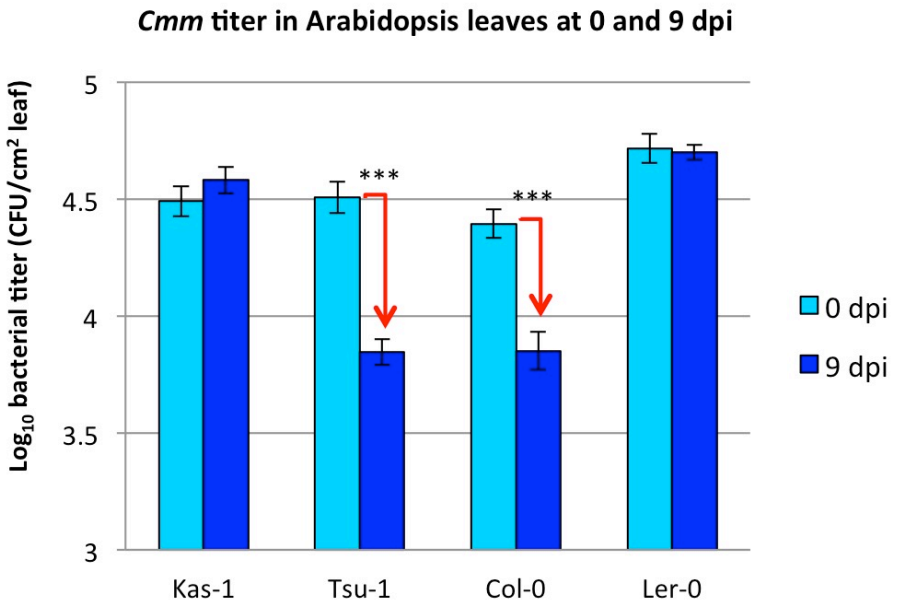


Figure 9. *In planta* survival of *C. michiganensis* subsp. *michiganensis* in four *A. thaliana* accessions. Fully expanded leaves of 30-day-old plants of *A. thaliana* accessions Kas-1, Tsu-1, Col-0 and Ler-0 were infiltrated with about 5×10^6 cfu/ml (OD_{600}=0.005) of *C. michiganensis* subsp. *michiganensis*. Bacteria were isolated from leaf discs collected at 0 and 9 dpi, and were plated on YGM agar supplemented with 50 µg/ml streptomycin. Bacterial titers were calculated using the results from four independent experiments each having four replicates for 0 dpi samples and six replicates for 9 dpi samples. The error bars represent standard error of the mean. Student's t-test was performed to compare the bacterial titers at 0 dpi and 9 dpi within each accession. ***, $p<0.001$.

References

Alarcon, C., Castro, J., Munoz, F., Arce-Johnson, P., and Delgado, J. (1998). Protein(s) from the Gram-positive bacterium *Clavibacter michiganensis* subsp. *michiganensis* induces a hypersensitive response in plants. Phytopathology **88,** 306-310.

Alfano, J.R., and Collmer, A. (2004). Type III secretion system effector proteins: double agents in bacterial disease and plant defense. Annu Rev Phytopathol **42,** 385-414.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A., and Zagnitko, O. (2008). The RAST Server: rapid annotations using subsystems technology. BMC Genomics **9,** 75.

Balaji, V., Sessa, G., and Smart, C.D. (2011). Silencing of host basal defense response-related gene expression increases susceptibility of *Nicotiana benthamiana* to *Clavibacter michiganensis* subsp. *michiganensis*. Phytopathology **101,** 349-357.

Balaji, V., Mayrose, M., Sherf, O., Jacob-Hirsch, J., Eichenlaub, R., Iraki, N., Manulis-Sasson, S., Rechavi, G., Barash, I., and Sessa, G. (2008). Tomato transcriptional changes in response to *Clavibacter michiganensis* subsp. *michiganensis* reveal a role for ethylene in disease development. Plant Physiol **146,** 1797-1809.

Bentley, S.D., Corton, C., Brown, S.E., Barron, A., Clark, L., Doggett, J., Harris, B., Ormond, D., Quail, M.A., May, G., Francis, D., Knudson, D., Parkhill, J., and Ishimaru, C.A. (2008). Genome of the actinomycete plant pathogen Clavibacter michiganensis subsp. sepedonicus suggests recent niche adaptation. J Bacteriol **190,** 2150-2160.

Bethke, G., Pecher, P., Eschen-Lippold, L., Tsuda, K., Katagiri, F., Glazebrook, J., Scheel, D., and Lee, J. (2012). Activation of the *Arabidopsis thaliana* mitogen-activated protein kinase MPK11 by the flagellin-derived elicitor peptide, flg22. Mol Plant Microbe Interact **25,** 471-480.

Biddle, J.A., Mcgee, D.C., and Braun, E.J. (1990). Seed transmission of *Clavibacter michiganense* subsp. nebraskense in corn. Plant Disease **74,** 908-911.

Block, A., Li, G., Fu, Z.Q., and Alfano, J.R. (2008). Phytopathogen type III effector weaponry and their plant targets. Curr Opin Plant Biol **11,** 396-403.

Boller, T., and Felix, G. (2009). A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. Annu Rev Plant Biol **60,** 379-406.

Bouwmeester, K., Han, M., Blanco-Portales, R., Song, W., Weide, R., Guo, L.-Y., van der Vossen, E.A.G., and Govers, F. (2014). The Arabidopsis lectin receptor kinase LecRK-I.9 enhances resistance to *Phytophthora infestans* in Solanaceous plants. Plant Biotechnol J **12,** 10-16.

Broman, K.W., Wu, H., Sen, S., and Churchill, G.A. (2003). R/qtl: QTL mapping in experimental crosses. Bioinformatics **19,** 889-890.

Burger, A., Grafen, I., Engemann, J., Niermann, E., Pieper, M., Kirchner, O., Gartemann, K.H., and Eichenlaub, R. (2005). Identification of homologues to the pathogenicity factor Pat-1, a putative serine protease of *Clavibacter michiganensis* subsp. *michiganensis*. Microbiol Res **160,** 417-427.

Caplan, J., Padmanabhan, M., and Dinesh-Kumar, S.P. (2008). Plant NB-LRR immune receptors: from recognition to transcriptional reprogramming. Cell Host Microbe **3,** 126-135.

Carver, T., Thomson, N., Bleasby, A., Berriman, M., and Parkhill, J. (2009). DNAPlotter: circular and linear interactive genome visualization. Bioinformatics **25,** 119-120.

Carver, T.J., Rutherford, K.M., Berriman, M., Rajandream, M.A., Barrell, B.G., and Parkhill, J. (2005). ACT: the Artemis Comparison Tool. Bioinformatics **21,** 3422-3423.

Casper-Lindley, C., Dahlbeck, D., Clark, E.T., and Staskawicz, B.J. (2002). Direct biochemical evidence for type III secretion-dependent translocation of the AvrBs2 effector protein into plant cells. Proc Natl Acad Sci U S A **99,** 8336-8341.

Chalupowicz, L., Cohen-Kandli, M., Dror, O., Eichenlaub, R., Gartemann, K.H., Sessa, G., Barash, I., and Manulis-Sasson, S. (2010). Sequential expression of bacterial virulence and plant defense genes during infection of tomato with *Clavibacter michiganensis* subsp. *michiganensis*. Phytopathology **100,** 252-261.

Chapman-Smith, A., and Cronan, J.E., Jr. (1999). Molecular biology of biotin attachment to proteins. J Nutr **129,** 477S-484S.

Chin, C.S., Alexander, D.H., Marks, P., Klammer, A.A., Drake, J., Heiner, C., Clum, A., Copeland, A., Huddleston, J., Eichler, E.E., Turner, S.W., and Korlach, J. (2013). Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods **10,** 563-569.

Chung, E.H., El-Kasmi, F., He, Y., Loehr, A., and Dangl, J.L. (2014). A plant phosphoswitch platform repeatedly targeted by type III effector proteins regulates the output of both tiers of plant immune receptors. Cell Host Microbe **16,** 484-494.

Coaker, G.L., Willard, B., Kinter, M., Stockinger, E.J., and Francis, D.M. (2004). Proteomic analysis of resistance mediated by Rcm 2.0 and Rcm 5.1, two loci controlling resistance to bacterial canker of tomato. Mol Plant Microbe Interact **17,** 1019-1028.

Contreras-Moreira, B., and Vinuesa, P. (2013). GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. Appl Environ Microbiol **79,** 7696-7701.

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

Darling, A.C., Mau, B., Blattner, F.R., and Perna, N.T. (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res **14,** 1394-1403.

De Boer, S.H., and Copeman, R.J. (1980). Bacterial ring rot testing with the indirect fluorescent antibody staining procedure. Am Potato J **57,** 457-465.

DeYoung, B.J., Qi, D., Kim, S.H., Burke, T.P., and Innes, R.W. (2012). Activation of a plant nucleotide binding-leucine rich repeat disease resistance protein by a modified self protein. Cell Microbiol **14,** 1071-1084.

Dong, X. (2004). NPR1, all things considered. Current Opinion in Plant Biology **7,** 547-552.

Dou, D., Kale, S.D., Wang, X., Jiang, R.H., Bruce, N.A., Arredondo, F.D., Zhang, X., and Tyler, B.M. (2008). RXLR-mediated entry of *Phytophthora sojae* effector Avr1b into soybean cells does not require pathogen-encoded machinery. Plant Cell **20,** 1930-1947.

Dreier, J., Meletzus, D., and Eichenlaub, R. (1997). Characterization of the plasmid encoded virulence region pat-1 of phytopathogenic *Clavibacter michiganensis* subsp. *michiganensis*. Mol Plant Microbe Interact **10,** 195-206.

Eichenlaub, R., and Gartemann, K.H. (2011). The *Clavibacter michiganensis* subspecies: molecular investigation of gram-positive bacterial plant pathogens. Annu Rev Phytopathol **49,** 445-464.

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman, B., Bibillo, A., Bjornson, K., Chaudhuri, B., Christians, F., Cicero, R., Clark, S., Dalal, R., Dewinter, A., Dixon, J., Foquet, M., Gaertner, A., Hardenbol, P., Heiner, C., Hester, K., Holden, D., Kearns, G., Kong, X., Kuse, R., Lacroix, Y., Lin, S., Lundquist, P., Ma, C., Marks, P., Maxham, M., Murphy, D., Park, I., Pham, T., Phillips, M., Roy, J., Sebra, R., Shen, G., Sorenson, J., Tomaney, A., Travers, K., Trulson, M., Vieceli, J., Wegener, J., Wu, D., Yang, A., Zaccarin, D., Zhao, P., Zhong, F., Korlach, J., and Turner, S. (2009). Real-time DNA sequencing from single polymerase molecules. Science **323,** 133-138.

Eitas, T.K., and Dangl, J.L. (2010). NB-LRR proteins: pairs, pieces, perception, partners, and pathways. Curr Opin Plant Biol **13,** 472-477.

Flor, H.H. (1971). Current Status of the Gene-For-Gene Concept. Annu Rev Phytopathol **9,** 275-296.

Flusberg, B.A., Webster, D.R., Lee, J.H., Travers, K.J., Olivares, E.C., Clark, T.A., Korlach, J., and Turner, S.W. (2010). Direct detection of DNA methylation during single-molecule, real-time sequencing. Nat Methods **7,** 461-465.

Francis, D.M., Kabelka, E., Bell, J., Franchino, B., and Clair, D.S. (2001). Resistance to bacterial canker in tomato (*Lycopersicon hirsutum* LA407) and its progeny derived from crosses to *L. esculentum*. Plant Dis **85,** 1171-1176.

Fu, Z.Q., Yan, S., Saleh, A., Wang, W., Ruble, J., Oka, N., Mohan, R., Spoel, S.H., Tada, Y., Zheng, N., and Dong, X. (2012). NPR3 and NPR4 are receptors for the immune signal salicylic acid in plants. Nature **486,** 228-232.

Gao, X., Chen, X., Lin, W., Chen, S., Lu, D., Niu, Y., Li, L., Cheng, C., McCormack, M., Sheen, J., Shan, L., and He, P. (2013). Bifurcation of Arabidopsis NLR immune signaling via Ca(2)(+)-dependent protein kinases. PLoS Pathog **9,** e1003127.

Gartemann, K.H., Abt, B., Bekel, T., Burger, A., Engemann, J., Flugel, M., Gaigalat, L., Goesmann, A., Grafen, I., Kalinowski, J., Kaup, O., Kirchner, O., Krause,

L., Linke, B., McHardy, A., Meyer, F., Pohle, S., Ruckert, C., Schneiker, S., Zellermann, E.M., Puhler, A., Eichenlaub, R., Kaiser, O., and Bartels, D. (2008). The genome sequence of the tomato-pathogenic actinomycete *Clavibacter michiganensis* subsp. *michiganensis* NCPPB382 reveals a large island involved in pathogenicity. J Bacteriol **190,** 2138-2149.

Glazebrook, J. (2005). Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. Annu Rev Phytopathol **43,** 205-227.

Gonzalez, A.J., and Trapiello, E. (2014). *Clavibacter michiganensis* subsp. *phaseoli* , a new subspecies pathogen in bean. Int J Syst Evol Microbiol **64,** 1752-1755.

Griesbach, E., Eisbein, K., and Sotirova, V. (2000). Induction of resistance to *Clavibacter michiganensis* subsp. *michiganensis*. Acta Physiol Plant **22,** 359-362.

Gross, D.C., Vidaver, A.K., and Keralis, M.B. (1979). Indigenous plasmids from phytopathogenic *Corynebacterium* species. J Gen Microbiol **115,** 479-489.

Holtsmark, I., Takle, G.W., and Brurberg, M.B. (2008). Expression of putative virulence factors in the potato pathogen *Clavibacter michiganensis* subsp. *sepedonicus* during infection. Arch Microbiol **189,** 131-139.

Jahr, H., Dreier, J., Meletzus, D., Bahro, R., and Eichenlaub, R. (2000). The endo-beta-1,4-glucanase CelA of *Clavibacter michiganensis* subsp. *michiganensis* is a pathogenicity determinant required for induction of bacterial wilt of tomato. Mol Plant Microbe Interact **13,** 703-714.

Jia, Y., McAdams, S.A., Bryan, G.T., Hershey, H.P., and Valent, B. (2000). Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. EMBO J **19,** 4004-4014.

Jiang, S., Yao, J., Ma, K.W., Zhou, H., Song, J., He, S.Y., and Ma, W. (2013). Bacterial effector activates jasmonate signaling by directly targeting JAZ transcriptional repressors. PLoS Pathog **9,** e1003715.

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

Katsir, L., Schilmiller, A.L., Staswick, P.E., He, S.Y., and Howe, G.A. (2008). COI1 is a critical component of a receptor for jasmonate and the bacterial virulence factor coronatine. Proc Natl Acad Sci U S A **105,** 7100-7105.

Ke, J., Khan, R., Johnson, T., Somers, D.A., and Das, A. (2001). High-efficiency gene transfer to recalcitrant plants by *Agrobacterium tumefaciens*. Plant Cell Rep **20,** 150-156.

Kenton, A., Parokonny, A.S., Gleba, Y.Y., and Bennett, M.D. (1993). Characterization of the *Nicotiana tabacum* L. genome by molecular cytogenetics. Mol Gen Genet **240,** 159-169.

Kim, Y., Tsuda, K., Igarashi, D., Hillmer, R.A., Sakakibara, H., Myers, C.L., and Katagiri, F. (2014). Mechanisms underlying robustness and tunability in a plant immune signaling network. Cell Host Microbe **15,** 84-94.

Kirchner, O., Gartemann, K.H., Zellermann, E.M., Eichenlaub, R., and Burger, A. (2001). A highly efficient transposon mutagenesis system for the tomato pathogen *Clavibacter michiganensis* subsp. *michiganensis*. Mol Plant Microbe Interact **14,** 1312-1318.

Koncz, C., and Schell, J. (1986). The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. Mol Gen Genet **204,** 383-396.

Konieczny, A., and Ausubel, F.M. (1993). A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. Plant J **4,** 403-410.

Kristensen, D.M., Kannan, L., Coleman, M.K., Wolf, Y.I., Sorokin, A., Koonin, E.V., and Mushegian, A. (2010). A low-polynomial algorithm for assembling clusters of orthologous groups from intergenomic symmetric best matches. Bioinformatics **26,** 1481-1487.

Laine, M.J., Nakhei, H., Dreier, J., Lehtila, K., Meletzus, D., Eichenlaub, R., and Metzler, M.C. (1996). Stable transformation of the gram-positive phytopathogenic bacterium *Clavibacter michiganensis* subsp. *sepedonicus* with several cloning vectors. Appl Environ Microbiol **62,** 1500-1506.

Laine, M.J., Haapalainen, M., Wahlroos, T., Kankare, K., Nissinen, R., Kassuwi, S., and Metzler, M.C. (2000). The cellulase encoded by the native plasmid of *Clavibacter michiganensis* ssp. *sepedonicus* plays a role in virulence and contains an expansin-like domain. Physiol Mol Plant Pathol **57,** 221-233.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods **9,** 357-359.

Lara-Ávila, J.P., Isordia-Jasso, M.I., Castillo-Collazo, R., Simpson, J., and Alpuche-Solís, Á.G. (2011). Gene expression analysis during interaction of tomato and related wild species with *Clavibacter michiganensis* subsp. *michiganensis*. Plant Mol Biol Report **30,** 498-511.

Larson, R.H. (1944). The ring rot bacterium in relation to tomato and eggplant. J Agric Res **69,** 309-325.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics **25,** 2078-2079.

Li, L., Stoeckert, C.J., Jr., and Roos, D.S. (2003). OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res **13,** 2178-2189.

Li, L., Li, M., Yu, L., Zhou, Z., Liang, X., Liu, Z., Cai, G., Gao, L., Zhang, X., Wang, Y., Chen, S., and Zhou, J.M. (2014). The FLS2-associated kinase BIK1 directly phosphorylates the NADPH oxidase RbohD to control plant immunity. Cell Host Microbe **15,** 329-338.

Li, X., Lin, H., Zhang, W., Zou, Y., Zhang, J., Tang, X., and Zhou, J.M. (2005). Flagellin induces innate immunity in nonhost interactions that is suppressed by *Pseudomonas syringae* effectors. Proc Natl Acad Sci U S A **102,** 12990-12995.

Lipka, V., Dittgen, J., Bednarek, P., Bhat, R., Wiermer, M., Stein, M., Landtag, J., Brandt, W., Rosahl, S., Scheel, D., Llorente, F., Molina, A., Parker, J., Somerville, S., and Schulze-Lefert, P. (2005). Pre- and postinvasion defenses both contribute to nonhost resistance in Arabidopsis. Science **310,** 1180-1183.

Lorang, J., Kidarsa, T., Bradford, C.S., Gilbert, B., Curtis, M., Tzeng, S.C., Maier, C.S., and Wolpert, T.J. (2012). Tricking the guard: exploiting plant defense for disease susceptibility. Science **338,** 659-662.

Lorenzo, O., Piqueras, R., Sanchez-Serrano, J.J., and Solano, R. (2003). ETHYLENE RESPONSE FACTOR1 integrates signals from ethylene and jasmonate pathways in plant defense. Plant Cell **15,** 165-178.

Lu, D., Wu, S., Gao, X., Zhang, Y., Shan, L., and He, P. (2010). A receptor-like cytoplasmic kinase, BIK1, associates with a flagellin receptor complex to initiate plant innate immunity. Proc Natl Acad Sci U S A **107,** 496-501.

Lu, F., Wang, H., Wang, S., Jiang, W., Shan, C., Li, B., Yang, J., Zhang, S., and Sun, W. (2014). Enhancement of innate immune system in monocot rice by transferring the dicotyledonous elongation factor Tu receptor EFR. J Integr Plant Biol **57**, 641-652.

Lu, M., Tang, X., and Zhou, J.M. (2001). Arabidopsis NHO1 is required for general resistance against *Pseudomonas* bacteria. Plant Cell **13,** 437-447.

Macho, A.P., Schwessinger, B., Ntoukakis, V., Brutus, A., Segonzac, C., Roy, S., Kadota, Y., Oh, M.H., Sklenar, J., Derbyshire, P., Lozano-Duran, R., Malinovsky, F.G., Monaghan, J., Menke, F.L., Huber, S.C., He, S.Y., and Zipfel, C. (2014). A bacterial tyrosine phosphatase inhibits plant pattern recognition receptor activation. Science **343,** 1509-1512.

Marefat, A., Ophel-Keller, K., and McKay, A. (2007). A real-time PCR assay for detection of *Clavibacter michiganensis* subsp. *insidiosus* in lucerne. Australas Plant Pathol **36,** 262-269.

Martin, G.B., Bogdanove, A.J., and Sessa, G. (2003). Understanding the functions of plant disease resistance proteins. Annu Rev Plant Biol **54,** 23-61.

McKay, J.K., Richards, J.H., Nemali, K.S., Sen, S., Mitchell-Olds, T., Boles, S., Stahl, E.A., Wayne, T., and Juenger, T.E. (2008). Genetics of drought adaptation in *Arabidopsis thaliana* II. QTL analysis of a new mapping population, KAS-1 x TSU-1. Evolution **62,** 3014-3026.

Meletzus, D., Bermphol, A., Dreier, J., and Eichenlaub, R. (1993). Evidence for plasmid-encoded virulence factors in the phytopathogenic bacterium *Clavibacter michiganensis* subsp. *michiganensis* NCPPB382. J Bacteriol **175,** 2131-2136.

Melotto, M., Underwood, W., Koczan, J., Nomura, K., and He, S.Y. (2006). Plant stomata function in innate immunity against bacterial invasion. Cell **126,** 969-980.

Meng, X., and Zhang, S. (2013). MAPK cascades in plant disease resistance signaling. Annu Rev Phytopathol **51,** 245-266.

Metzler, M.C., Laine, M.J., and De Boer, S.H. (1997). The status of molecular biological research on the plant pathogenic genus *Clavibacter*. FEMS Microbiol Lett **150,** 1-8.

Michalke, A., Galla, H.J., and Steinem, C. (2001). Channel activity of a phytotoxin of *Clavibacter michiganense* ssp. *nebraskense* in tethered membranes. Eur Biophys J **30,** 421-429.

Mysore, K.S., and Ryu, C.-M. (2004). Nonhost resistance: how much do we know? Trends Plant Sci **9,** 97-104.

Nakamura, K., Oshima, T., Morimoto, T., Ikeda, S., Yoshikawa, H., Shiwa, Y., Ishikawa, S., Linak, M.C., Hirai, A., Takahashi, H., Altaf-Ul-Amin, M., Ogasawara, N., and Kanaya, S. (2011). Sequence-specific error profile of Illumina sequencers. Nucleic Acids Res **39,** e90-e90.

NÉMeth, J., Laszlo, E., and EmÖDy, L. (1991). *Clavibacter michiganensis* ssp. *insidiosus* in lucerne seeds. EPPO Bulletin **21,** 713-718.

Nissinen, R., Xia, Y., Mattinen, L., Ishimaru, C.A., Knudson, D.L., Knudson, S.E., Metzler, M., and Pirhonen, M. (2009). The putative secreted serine protease Chp-7 is required for full virulence and induction of a nonhost hypersensitive response by *Clavibacter michiganensis* subsp. *sepedonicus*. Mol Plant Microbe Interact **22,** 809-819.

Nissinen, R., Lai, F.M., Laine, M.J., Bauer, P.J., Reilley, A.A., Li, X., De Boer, S.H., Ishimaru, C.A., and Metzler, M.C. (1997). *Clavibacter michiganensis* subsp. *sepedonicus* elicits a hypersensitive response in tobacco and secretes hypersensitive response-inducing protein(s). Phytopathology **87,** 678-684.

Nurnberger, T., and Lipka, V. (2005). Non-host resistance in plants: new insights into an old phenomenon. Mol Plant Pathol **6,** 335-345.

Papadopoulou, K., Melton, R.E., Leggett, M., Daniels, M.J., and Osbourn, A.E. (1999). Compromised disease resistance in saponin-deficient plants. Proc Natl Acad Sci U S A **96,** 12923-12928.

Paschke, M.C., and Van Alfen, N.K. (1993). Extracellular polysaccharide impaired mutants of *Clavibacter michiganense* subsp. *insidiosum*. Physiol Mol Plant Pathol **42,** 309-319.

Pre, M., Atallah, M., Champion, A., De Vos, M., Pieterse, C.M., and Memelink, J. (2008). The AP2/ERF domain transcription factor ORA59 integrates jasmonic acid and ethylene signals in plant defense. Plant Physiol **147,** 1347-1357.

Qi, Y., and Katagiri, F. (2009). Purification of low-abundance Arabidopsis plasma-membrane protein complexes and identification of candidate components. Plant J **57,** 932-944.

Quail, M., Smith, M., Coupland, P., Otto, T., Harris, S., Connor, T., Bertoni, A., Swerdlow, H., and Gu, Y. (2012). A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. BMC Genomics **13,** 341.

Reverchon, S., Rouanet, C., Expert, D., and Nasser, W. (2002). Characterization of indigoidine biosynthetic genes in *Erwinia chrysanthemi* and role of this blue pigment in pathogenicity. J Bacteriol **184,** 654-665.

Ries, S.M., and Strobel, G.A. (1972). Biological properties and pathological role of a phytotoxic glycopeptide from *Corynebacterium insidiosum*. Physiol Plant Pathol **2,** 133-142.

Roberts, R., Carneiro, M., and Schatz, M. (2013). The advantages of SMRT sequencing. Genome Biol **14,** 405.

Samac, D.A., Nix, R.J., and Oleson, A.E. (1998). Transmission frequency of *Clavibacter michiganensis* subsp. *insidiosus* to alfalfa seed and identification of the bacterium by PCR. Plant Dis **82,** 1362-1367.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A **74,** 5463-5467.

Savidor, A., Chalupowicz, L., Teper, D., Gartemann, K.H., Eichenlaub, R., Manulis-Sasson, S., Barash, I., and Sessa, G. (2014). *Clavibacter michiganensis* subsp. *michiganensis* Vatr1 and Vatr2 transcriptional regulators are required for virulence in tomato. Mol Plant Microbe Interact **27,** 1035-1047.

Savidor, A., Teper, D., Gartemann, K.H., Eichenlaub, R., Chalupowicz, L., Manulis-Sasson, S., Barash, I., Tews, H., Mayer, K., Giannone, R.J., Hettich, R.L., and Sessa, G. (2012). The *Clavibacter michiganensis* subsp. *michiganensis*-tomato interactome reveals the perception of pathogen by the host and suggests mechanisms of infection. J Proteome Res **11,** 736-750.

Schechter, L.M., Roberts, K.A., Jamir, Y., Alfano, J.R., and Collmer, A. (2004). *Pseudomonas syringae* type III secretion system targeting signals and novel effectors studied with a Cya translocation reporter. J Bacteriol **186,** 543-555.

Scheible, W.R., Fry, B., Kochevenko, A., Schindelasch, D., Zimmerli, L., Somerville, S., Loria, R., and Somerville, C.R. (2003). An Arabidopsis mutant resistant to thaxtomin A, a cellulose synthesis inhibitor from *Streptomyces* species. Plant Cell **15,** 1781-1794.

Schoonbeek, H.-j., Wang, H.-H., Stefanato, F.L., Craze, M., Bowden, S., Wallington, E., Zipfel, C., and Ridout, C.J. (2015). Arabidopsis EF-Tu receptor enhances bacterial disease resistance in transgenic wheat. New Phytol **206,** 606-613.

Schurholz, T., Wilimzig, M., Katsiou, E., and Eichenlaub, R. (1991). Anion channel forming activity from the plant pathogenic bacterium *Clavibacter michiganense* ssp. *nebraskense*. J Membr Biol **123,** 1-8.

Seyfferth, C., and Tsuda, K. (2014). Salicylic acid signal transduction: the initiation of biosynthesis, perception and transcriptional reprogramming. Front Plant Sci **5,** 697.

Shao, F., Golstein, C., Ade, J., Stoutemyer, M., Dixon, J.E., and Innes, R.W. (2003). Cleavage of Arabidopsis PBS1 by a bacterial type III effector. Science **301,** 1230-1233.

Sheard, L.B., Tan, X., Mao, H., Withers, J., Ben-Nissan, G., Hinds, T.R., Kobayashi, Y., Hsu, F.F., Sharon, M., Browse, J., He, S.Y., Rizo, J., Howe, G.A., and Zheng, N. (2010). Jasmonate perception by inositol-phosphate-potentiated COI1-JAZ co-receptor. Nature **468,** 400-405.

Sohn, K.H., Saucet, S.B., Clarke, C.R., Vinatzer, B.A., O'Brien, H.E., Guttman, D.S., and Jones, J.D. (2012). HopAS1 recognition significantly contributes to Arabidopsis nonhost resistance to *Pseudomonas syringae* pathogens. New Phytol **193,** 58-66.

Sonnhammer, E.L., and Durbin, R. (1995). A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. Gene **167,** GC1-10.

Sory, M.P., and Cornelis, G.R. (1994). Translocation of a hybrid YopE-adenylate cyclase from *Yersinia enterocolitica* into HeLa cells. Mol Microbiol **14,** 583-594.

Spoel, S.H., and Dong, X. (2008). Making sense of hormone crosstalk during plant immune responses. Cell Host Microbe **3,** 348-351.
Starr, M. (1958). The blue pigment of *Corynebacterium insidiosum*. Arch Mikrobiol **30,** 325-334.
Stergiopoulos, I., and de Wit, P.J. (2009). Fungal effector proteins. Annu Rev Phytopathol **47,** 233-263.
Stork, I., Gartemann, K.H., Burger, A., and Eichenlaub, R. (2008). A family of serine proteases of *Clavibacter michiganensis* subsp. *michiganensis*: *chpC* plays a role in colonization of the host plant tomato. Mol Plant Pathol **9,** 599-608.
Stotz, H.U., Mitrousia, G.K., de Wit, P.J., and Fitt, B.D. (2014). Effector-triggered defence against apoplastic fungal pathogens. Trends Plant Sci **19,** 491-500.
Strobel, G.A. (1977). Bacterial phytotoxins. Annu Rev Microbiol **31,** 205-224.
Strobel, G.A., and Hess, W.M. (1968). Biological Activity of a Phytotoxic Glycopeptide Produced by *Corynebacterium sepedonicum*. Plant Physiol **43,** 1673-1688.
Takahashi, H., Kumagai, T., Kitani, K., Mori, M., Matoba, Y., and Sugiyama, M. (2007). Cloning and characterization of a *Streptomyces* single module type non-ribosomal peptide synthetase catalyzing a blue pigment synthesis. J Biol Chem **282,** 9073-9081.
Tatusov, R.L., Galperin, M.Y., Natale, D.A., and Koonin, E.V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res **28,** 33-36.
Thaler, J.S., Humphrey, P.T., and Whiteman, N.K. (2012). Evolution of jasmonate and salicylate signal crosstalk. Trends Plant Sci **17,** 260-270.
Thines, B., Katsir, L., Melotto, M., Niu, Y., Mandaokar, A., Liu, G., Nomura, K., He, S.Y., Howe, G.A., and Browse, J. (2007). JAZ repressor proteins are targets of the SCF(COI1) complex during jasmonate signalling. Nature **448,** 661-665.
Thomas, W.J., Thireault, C.A., Kimbrel, J.A., and Chang, J.H. (2009). Recombineering and stable integration of the *Pseudomonas syringae* pv. syringae 61 hrp/hrc cluster into the genome of the soil bacterium *Pseudomonas fluorescens* Pf0-1. Plant J **60,** 919-928.
Tripathi, J.N., Lorenzen, J., Bahar, O., Ronald, P., and Tripathi, L. (2014). Transgenic expression of the rice Xa21 pattern-recognition receptor in banana (Musa sp.) confers resistance to *Xanthomonas campestris* pv. musacearum. Plant Biotechnol J**12,** 663-673.
Tsiantos, J. (1987). Transmission of bacterium *Corynebacterium michiganense* pv. *michiganense* by seeds. J Phytopathol **119,** 142-146.
Tsuda, K., Sato, M., Stoddard, T., Glazebrook, J., and Katagiri, F. (2009). Network properties of robust immunity in plants. PLoS Genet **5,** e1000772.
Tsuda, K., Mine, A., Bethke, G., Igarashi, D., Botanga, C.J., Tsuda, Y., Glazebrook, J., Sato, M., and Katagiri, F. (2013). Dual regulation of gene expression mediated by extended MAPK activation and salicylic acid contributes to robust innate immunity in *Arabidopsis thaliana*. PLoS Genet **9,** e1004015.
Van Alfen, N.K., and Turner, N.C. (1975). Changes in alfalfa stem conductance induced by *Corynebacterium insidiosum* toxin. Plant Physiol **55,** 559-561.

Van Den Bulk, R.W., Löffler, H.J.M., and Dons, J.J.M. (1989). Effect of phytotoxic compounds produced by *Clavibacter michiganensis* subsp. *michiganensis* on resistant and susceptible tomato plants. Eur J Plant Pathol **95,** 107-117.

van Ooijen, G., van den Burg, H.A., Cornelissen, B.J., and Takken, F.L. (2007). Structure and function of resistance proteins in solanaceous plants. Annu Rev Phytopathol **45,** 43-72.

Vancanneyt, G., Schmidt, R., O'Connor-Sanchez, A., Willmitzer, L., and Rocha-Sosa, M. (1990). Construction of an intron-containing marker gene: splicing of the intron in transgenic plants and its use in monitoring early events in *Agrobacterium*-mediated plant transformation. Mol Gen Genet **220,** 245-250.

Vlot, A.C., Dempsey, D.A., and Klessig, D.F. (2009). Salicylic Acid, a multifaceted hormone to combat disease. Annu Rev Phytopathol **47,** 177-206.

Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C.A., Zeng, Q., Wortman, J., Young, S.K., and Earl, A.M. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One **9,** e112963.

Wang, K., Senthil-Kumar, M., Ryu, C.M., Kang, L., and Mysore, K.S. (2012). Phytosterols play a key role in plant innate immunity against bacterial pathogens by regulating nutrient efflux into the apoplast. Plant Physiol **158,** 1789-1802.

Wang, L., Tsuda, K., Sato, M., Cohen, J.D., Katagiri, F., and Glazebrook, J. (2009). Arabidopsis CaM binding protein CBP60g contributes to MAMP-induced SA accumulation and is involved in disease resistance against Pseudomonas syringae. PLoS Pathog **5,** e1000301.

Wang, L., Tsuda, K., Truman, W., Sato, M., Nguyen le, V., Katagiri, F., and Glazebrook, J. (2011). CBP60g and SARD1 play partially redundant critical roles in salicylic acid signaling. Plant J **67,** 1029-1041.

Whisson, S.C., Boevink, P.C., Moleleki, L., Avrova, A.O., Morales, J.G., Gilroy, E.M., Armstrong, M.R., Grouffaud, S., van West, P., Chapman, S., Hein, I., Toth, I.K., Pritchard, L., and Birch, P.R. (2007). A translocation signal for delivery of oomycete effector proteins into host plant cells. Nature **450,** 115-118.

Willats, W.G.T., Orfila, C., Limberg, G., Buchholt, H.C., van Alebeek, G.-J.W.M., Voragen, A.G.J., Marcus, S.E., Christensen, T.M.I.E., Mikkelsen, J.D., Murray, B.S., and Knox, J.P. (2001). Modulation of the degree and pattern of methyl-esterification of pectic homogalacturonan in plant cell walls: implications for pectin methyl esterase action, matrix properties, and cell adhesion. J Biol Chem **276,** 19404-19413.

Wilson, K. (2001). Preparation of Genomic DNA from Bacteria. Curr Protoc Mol Biol, **Unit 2.4**.

Xiang, T., Zong, N., Zou, Y., Wu, Y., Zhang, J., Xing, W., Li, Y., Tang, X., Zhu, L., Chai, J., and Zhou, J.M. (2008). *Pseudomonas syringae* effector AvrPto blocks innate immunity by targeting receptor kinases. Curr Biol **18,** 74-80.

Xie, D.X., Feys, B.F., James, S., Nieto-Rostro, M., and Turner, J.G. (1998). COI1: an Arabidopsis gene required for jasmonate-regulated defense and fertility. Science **280,** 1091-1094.

Xin, X.F., and He, S.Y. (2013). *Pseudomonas syringae* pv. *tomato* DC3000: a model pathogen for probing disease susceptibility and hormone signaling in plants. Annu Rev Phytopathol **51,** 473-498.

Yasuhara-Bell, J., and Alvarez, A.M. (2015). Seed-associated subspecies of the genus *Clavibacter* are clearly distinguishable from *Clavibacter michiganensis* subsp. *michiganensis*. Int J Syst Evol Microbiol **65,** 811-826.

Yim, K.-O., Lee, H.-I., Kim, J.-H., Lee, S.-D., Cho, J.-H., and Cha, J.-S. (2011). Characterization of phenotypic variants of *Clavibacter michiganensis* subsp. michiganensis isolated from *Capsicum annuum*. Eur J Plant Pathol **133,** 559-575.

Yu, D., Xu, F., Valiente, J., Wang, S., and Zhan, J. (2013). An indigoidine biosynthetic gene cluster from *Streptomyces chromofuscus* ATCC 49982 contains an unusual IndB homologue. J Ind Microbiol Biotechnol **40,** 159-168.

Zaluga, J., Van Vaerenbergh, J., Stragier, P., Maes, M., and De Vos, P. (2013). Genetic diversity of non-pathogenic *Clavibacter* strains isolated from tomato seeds. Syst Appl Microbiol **36,** 426-435.

Zaluga, J., Stragier, P., Baeyen, S., Haegeman, A., Van Vaerenbergh, J., Maes, M., and De Vos, P. (2014). Comparative genome analysis of pathogenic and non-pathogenic *Clavibacter* strains reveals adaptations to their lifestyle. BMC Genomics **15,** 392.

Zhang, J., Shao, F., Li, Y., Cui, H., Chen, L., Li, H., Zou, Y., Long, C., Lan, L., Chai, J., Chen, S., Tang, X., and Zhou, J.M. (2007). A *Pseudomonas syringae* effector inactivates MAPKs to suppress PAMP-induced immunity in plants. Cell Host Microbe **1,** 175-185.

Zhang, L., Du, L., Shen, C., Yang, Y., and Poovaiah, B.W. (2014). Regulation of plant immunity through ubiquitin-mediated modulation of Ca(2+) -calmodulin-AtSR1/CAMTA3 signaling. Plant J **78,** 269-281.

Zhang, Z., Wu, Y., Gao, M., Zhang, J., Kong, Q., Liu, Y., Ba, H., Zhou, J., and Zhang, Y. (2012). Disruption of PAMP-induced MAP kinase cascade by a *Pseudomonas syringae* effector activates plant immunity mediated by the NB-LRR protein SUMM2. Cell Host Microbe **11,** 253-263.

Zheng, X.Y., Spivey, N.W., Zeng, W., Liu, P.P., Fu, Z.Q., Klessig, D.F., He, S.Y., and Dong, X. (2012). Coronatine promotes Pseudomonas syringae virulence in plants by activating a signaling cascade that inhibits salicylic acid accumulation. Cell Host Microbe **11,** 587-596.

Zipfel, C. (2014). Plant pattern-recognition receptors. Trends Immunol **35,** 345-351.