Comparative genomics to develop a specific multiplex PCR assay for detection of Clavibacter michiganensis

Shree P. Thapa, Michael O'Leary, Marie Agnes Jacques, Robert Gilbertson, Gitta Coaker

TITLE: Comparative genomics to develop a specific multiplex PCR assay for detection of *Clavibacter michiganensis*

AUTHORS:

Authors: Shree P. Thapa[1], Michael O'Leary[1], Marie-Agnès Jacques[2], Robert L. Gilbertson[1], Gitta Coaker[1*]

AFFILIATIONS:

[1]Department of Plant Pathology, University of California, Davis, California, United States of America

[2]IRHS, Agrocampus-Ouest, INRA, Université d'Angers, SFR 4207 Quasav, Beaucouzé, France

*Corresponding Author: E-mail: glcoaker@ucdavis.edu; Tel. +1 530-752-6541; Fax. +1 530-752-5674

RUNNING TITLE: Comparative genomics based *Clavibacter* detection

Keywords: Multiplex PCR, diagnostic, *Clavibacter michiganensis*, comparative genomics, tomato

Funding: GC, SP, and RG were supported by grants from the California Tomato Research Institute awarded to GC and RG (091). GC and SP were supported by a grant from USDA-NIFA (2015-67013-23082, awarded to GC) and grants awarded by the

California Tomato Research Institute and the California Department of Food and Agriculture (17-0275-047-SC).

ABSTRACT

Clavibacter michiganensis (*Cm*) is a Gram-positive bacterial pathogen that proliferates in the xylem vessels of tomato, causing bacterial wilt and canker symptoms. Accurate detection is a crucial step in confirming outbreaks of bacterial canker and developing management strategies. A major problem with existing detection methods are false positive and negative results. Here, we report the use of comparative genomics of 37 diverse *Clavibacter* strains, including 21 strains sequenced in this study, to identify specific sequences that are *Cm* detection targets. Genome-wide phylogenic analyses revealed additional diversity within the genus *Clavibacter*. Pathogenic *Cm* strains varied in plasmid composition, highlighting the need for detection methods based on chromosomal targets. We utilized sequences of *Cm* specific loci two develop a multiplex PCR based diagnostic platform using two *Cm* chromosomal genes (*rhuM* and *tomA*) and an internal control amplifying both bacterial and plant DNA (16s rRNA). The multiplex PCR assay specifically detected *Cm* strains from a panel of 110 additional bacteria, including other *Clavibacter* species and bacterial pathogens of tomato. The assay was adapted to detect the presence of *Cm* in seeds and tomato plant materials with high sensitivity and specificity. In conclusion, the described method represents a robust, specific tool for detection of *Cm* in tomato seeds and infected plants.

INTRODUCTION

Clavibacter species are xylem colonizing Gram-positive bacteria that can infect both monocots and dicots. *Clavibacter* subspecies have been recently elevated to the species level based on average nucleotide identities derived from whole genome analyses and digital DNA-DNA hybridization (Li et al. 2018). *Clavibacter michiganensis* (*Cm*) causes bacterial canker and is one of the most important bacterial diseases of tomato (de León et al. 2011; Mansfield et al. 2012). *Cm* gains entry into the plant through contaminated seeds, wounds, broken trichomes, or hydathodes, and then proliferates in xylem vessels (Eichenlaub et al. 2006; Sen et al. 2014). Systemic infection results in unilateral leaf wilting, necrosis of the entire plant, necrosis of stems and petioles as well as formation of cankers (de León et al. 2011). *Cm* can also form localized infections resulting in fruit spotting “bird’s eye” and blister-like spots on vegetative tissues (de León et al. 2011).

Bacterial canker can cause substantial losses in greenhouse and field production systems. Since the first report of the disease in Michigan in 1910, it has spread throughout the world (Smith, 1910). The economic threat and the challenges to manage this disease have led to the inclusion of this pathogen as a quarantine agent in Europe, Asia, Africa and the Caribbean (Anonymous, 2000; Eichenlaub & Gartemann, 2011). Despite the importance of this disease, there currently are no resistant commercial cultivars or effective chemical control strategies. At present, control and management of

this disease rely primarily on the use of clean seed, healthy transplant practices, sanitation and crop rotation.

Effective disease management also requires precise, reliable and sensitive detection methods. Historically, methods are based on semi-selective media followed by pathogenicity tests (Fatmi and Schad, 2002; Fatmi et al. 2017). These diagnostic assays are expensive, require laboratory and greenhouse facilities and are unsuitable for large-scale screening. Current *Cm* detection strategies mainly utilize serological (Franken et al. 1993) or PCR-based methods with primers targeting different loci, oftenfollowed by bacterial isolation on semi-selective media (Dreier et al. 1995; Louws et al. 1998; Kaneshiro et al. 2006; Milijašević-Marčić et al. 2012). However, the current serological techniques such as ELISA and immunostrips are not reliable due to their cross reaction with other *Clavibacter* species (Franken et al. 1993; Dreier et al. 1995; Jacques et al. 2012).

The presence of closely related *Clavibacter* endophytes in tomato complicates precise detection of *Cm*, making it difficult to differentiate pathogenic *Cm* from other *Clavibacter* species. Non-pathogenic *Clavibacter* has been frequently found associated with tomato plants and seed (Jacques et al. 2012; Zaluga et al. 2013; Yasuhara-Bell & Alvarez, 2015; Thapa et al. 2017). Furthermore, analyses of *Cm* strains collected from diverse geographical origins have revealed considerable genetic diversity, particularly with respect to plasmid composition, plasmid number, and the presence or absence of putative chromosomally encoded pathogenicity genes (Jacques et al. 2012; Milijašević-

Marčić et al. 2012; Tancos et al. 2015; Thapa et al. 2017). The presence of endophytic *Clavibacter* species in tomato and the diversity in *Cm* highlights the need for a rapid, specific and accurate method for *Cm* from seed, plants and other substrates.

PCR is rapid, highly sensitive and can be specific for identification of bacterial pathogens (Schaad & Frederick, 2002; Vincelli & Tisserat, 2008; McNally et al. 2016). Several pairs of primers have been designed for detecting *Cm* (Alvarez, 2005; Kaneshiro et al. 2006; Kleitman et al. 2008; Santos et al.1997). However, most of these primers target genes present in plasmids (pCM1 and pCM2) and the *chp/tomA* pathogenicity island based on the reference strain NCPPB382 (Alvarez, 2005; Kaneshiro et al. 2006; Kleitman et al. 2008). False positive and negative results have been reported with most primers, and different *Cm* strains can possess variable plasmid components yet retain similar virulence (Louws et al. 1998; Alvarez, 2005; Kleitman et al. 2008; Tancos et al. 2015; Thapa et al. 2017). Therefore, there is a need for a robust diagnostic assay for *Cm* capable of targeting multiple genes that can differentiate *Cm* from other *Clavibacter* species and bacterial pathogens associated with tomato.

Multiplex PCR (mPCR) can be used to amplify multiple genetic loci, facilitating detection of several targets in a single pathogen (Bertolini et al. 2003; Ozdemir, 2009; Thapa et al. 2013; Thapa et al. 2012). Here, we used comparative genomics to analyze a wide range of strains representing the genus *Clavibacter*. This analysis revealed diversity among *Cm* strains, including multiple subgroups and variable plasmid components. Genome analyses were used to identify highly conserved multiplex PCR targets within

Cm strains. We have also validated the specificity and sensitivity of this assay on vegetative and seed samples using diverse *Cm* strains including those collected from a recent outbreak in California.

MATERIALS AND METHODS

Bacterial strains, DNA extraction and plasmids. The bacterial strains used in this study are listed in Table S1. *Cm* strains were grown in tryptone broth with yeast (TBY) and on D2 semi-selective media at 28°C (Kirchner et al. 2001). For DNA extraction, overnight cultures were grown in TBY broth at 28°C on a rotary shaker at 200 rpm. DNA was extracted from infected plant materials or bacterial cultures for PCR using either a crude lysis method or CTAB purification (Murray & Thompson, 1980). For the lysis method, 100mg of plant material or bacterial cells were resuspended in extraction buffer (7.75g/l Na_2HPO_4, 1.65g/l KH_2PO_4, 0.2ml/l Tween 20). The supernatant was then added to cell lysis buffer (3% SDS, 50µM Tris, and 10% sucrose) and incubated at 100^0C in heat block for 15 min. The samples were centrifuged at 11,000 g for 5 min. The supernatant was transferred into new Eppendorf tube. DNA was precipitated with 2x volume of 100% ethanol, centrifuged at 11,000 g for 5 min, and the supernatant discarded. The pellet was then washed with 70% ethanol, air dried and mixed in dH_2O, and used for multiplex PCR (mPCR) analyses. CTAB DNA extraction was performed as previously described (Murray & Thompson, 1980).

Genome sequencing, assembly and annotation. *Cm* strains sequenced in the present study are shown in Table 1. Total genomic DNA was isolated with the Promega Wizard® Genomic DNA Purification Kit. For library construction for the Illumina platform, genomic DNA was sheared with a Covaris E220 sonicator, followed by end repair and phosphorylation. Fragmented DNA was used to construct paired-end libraries. Libraries

were checked for size with a High Sensitivity DNA Kit (Agilent) in the Bioanalyzer 2100, with an expected average size of 250 bp. Libraries were pooled and sequenced with paired-end reads on the HiSeq 2500 system (Illumina). Genome sequencing was performed at the UC Davis DNA Technologies Core Facility. Genomes sequenced on Illumina were *de novo* assembled with the SPAdes (Bankevich et al. 2012). The draft genomes were annotated with Prokka and the NCBI Prokaryotic Genome Annotation Pipeline (Seemann, 2014; Tatusova et al. 2016).

Phylogenetic Analyses. Orthologous genes of *Clavibacter* strains were predicted with the OrthoMCL pipeline (Li et al. 2003). All-versus-all BLASTN (*E* value < 1×10^{-5}, alignment coverage > 50%) comparison of all gene sequences for each species was performed, and orthologous genes were clustered with OrthoMCL (Li et al. 2003). The normalized scores were fed into the MCL algorithm to classify the genes into predicted orthologous gene families with a default inflation parameter of 1.5. We performed multiple alignments of gene sequences with Prank (Loytynoja & Goldman, 2005). Alignments were concatenated by FASconCAT yielding a gene supermatrix (Kuck & Meusemann, 2010). A maximum likelihood approach was used to reconstruct the phylogenetic tree using RAxML software (Stamatakis, 2006). Bootstrapping was performed with 1000 replicates. The resulting phylogeny was visualized with FigTree (Rambaut, 2012).

Comparative genomics and primer design. Comparative analysis was performed using 145,929 genes from 21 sequenced genomes the present study and 24 published

Clavibacter genomes. In order to identify gene conservation across *Cm*, we performed whole genome clustering analyses based on sequence identity using CD-HIT software with an identity threshold set to 95% (Li & Godzik, 2006). Primers were designed based upon conserved sequences that were present only in *Cm* and absent in other *Clavibacter* or bacterial strains. Primers were designed based on specificity, amplicon size, and compatibility in mPCR using the Primer3 (http://bioinfo.ut.ee/primer3-0.4.0/). Primer design was optimized to minimize secondary structure, to have compatible annealing temperature, and minimal cross dimerization. Primer specificity was tested in silico by a similarity search against NCBI sequence database (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Three primer sets, two *Cm* specific and one internal control, were designed to amplify RhuM (CMM_2694), TomA (CMM_0090) and 16S (CMM_RNA_0001) (Table S2).

Multiplex PCR (mPCR) assay and limit of detection. Multiplex PCR was performed with primers for RhuM-F/R, TomA-F1/R1 and 16sR-F/R primers. The specificity of the mPCR assay was determined by using 75 *Cm,* 11 *Clavibacter sepedonicus*, 1 *Clavibacter insidious*, 2 *Clavibacter tessellarius*, 6 endophytic *Clavibacter*, and 15 other bacterial pathogens of tomato (Table S1). Total genomic DNA was isolated with the Promega Wizard® Genomic DNA Purification Kit. Each primer combination was tested using the C1000 Touch™ Thermal Cycler (Bio-Rad, Hercules, CA, USA). mPCR was performed in 25 µl final volume with 1-4 µl DNA, 0.7µM (RhuM-F/R), 0.5µM (TomA-F1/R1), 0.3µM (16sR-F/R) primers, 4 µl GC buffer and 0.02 U/µl iProof polymerase (Bio-Rad). A temperature gradient of 58 to 65°C was used to determine the optimum

annealing temperature. All PCRs amplifications were carried out in a C1000 Touch™ Thermal Cycler (Bio-Rad). Final PCR conditions were as follows: initial denaturation of 98°C for 2 min, followed by 30 cycles at 98°C for 30 s, 60°C for 30 s, 68°C for 1 min, and a final extension step of 68°C for 7 min. mPCR products were electrophoresed for 30 min at 130 V on a 1% agarose gels.

To evaluate the sensitivity of the mPCR assay, bacterial cultures and genomic DNA of *Cm* CASJ002 were serially diluted and used as a template. For DNA, 10 ng of total *Cm* genomic DNA was diluted and used for mPCR assay in a C1000 Touch™ Thermal Cycler (Bio-Rad) following conditions as above. For bacterial cultures, *Cm* (CASJ002) was grown on TBY broth and cells were serially diluted followed by DNA extraction using the lysis method. Each of the dilutions (100 µl) was spread on TBY and incubated for 48 h at 28°C to determine the concentration of cells in the suspension.

Plant growth and pathogenicity assays. Tomato plants cv. Early Pak No. 7 (EP-7) were grown in the greenhouse at 23°C, 50% relative humidity with a 14 h photoperiod and supplemental lighting as necessary (Thapa et al. 2015). Three to four-week-old tomato plants were used for inoculation assays. Tomato plants were inoculated with *Cm* strain CASJ002 as previously described (Thapa et al. 2017). To obtain DNA from inoculated plants, samples were collected 7 and 14 days post-inoculation. One cm stem segments were cut at 1, 3, 6, 9, 15 cm above the inoculation site and from apical regions. Stem segments were surface disinfected in 75% ethanol, the outer epidermis removed, and the remaining segment was ground in 500 µl dH_2O. Then, 100 µl of the

resulting supernatant was used for DNA extraction using CTAB, and 100 µl was used for determination of bacterial titers.

Seed infestation, sampling and DNA extraction. For seed contamination, CASJ002 inoculum was prepared as described above. Inoculum was prepared by diluting cells in 10 mM $MgCl_2$ to a concentration of approximately 10^7 CFU/ml. EP-7 Seed lots (10 g) were then incubated with 40ml of *Cm* CASJ002 in H_2O for 24 hrs at 37°C. After 24h, the liquid was removed and seeds were dried at 24C on filter paper for 2 days and kept for 2-5 weeks at room temperature. The detection threshold was evaluated by adding different numbers of contaminated seeds to clean seed samples of 10,000 seeds. Samples of 10,000 (~30 g) seeds were suspended in 120 ml of sterile PBS buffer. The samples were incubated at 4°C for at least 12 h, shaken for 2–3 h at room temperature, and then transferred to whirl pack bags (Whirl-pak Nasco, Modesto, CA) and processed in a stomacher (Stomacher 400, Seward, England) for 20 min. Both the lysis and CTAB procedure were used for the isolation of DNA to be used for mPCR. The PCR conditions used were as described above.

Collection of field samples. During 2016, 21 samples of field-grown tomato plants exhibiting bacterial canker-like symptoms were collected from the California Central Valley (Table S3). Stem and leaf samples of plants exhibiting disease symptoms were used for DNA extraction and bacterial isolation. DNA was extracted from 1 mg plant materials using both lysis and CTAB procedures and used in mPCR tests. *Cm* isolation was performed using the semi-selective D2 media (Kado & Heskett, 1970).

233

234 **Accession numbers.** Whole genome sequences of strains sequenced in this study

235 were submitted to NCBI under the accession numbers listed in Table 1.

236

RESULTS

***C. michiganensis* genome sequencing and phylogenomics.** We sequenced 21 *Cm* genomes from strains isolated from different locations world-wide using the Illumina platform (Fig. S1A). These 21 *Cm* strains were verified to be pathogenic on tomato after inoculation of the cultivar Early Pak 7 and are able to induce leaflet necrosis, wilting and stem canker symptoms (Fig. 1). Genomes were sequenced on the Illumina platform, *de novo* assembled with the SPAdes, and annotated with Prokka and the NCBI Prokaryotic Genome Annotation Pipeline (Table 1). To facilitate a high-resolution comparison of *Clavibacter* species, the core genome of all 21 strains sequenced in the present study and 16 other published genomes, including five *Clavibacter* species pathogenic on other crops and five tomato endophytes, were analyzed. Using the orthoMCL pipeline, we identified a total of 1,258 orthologous genes that are conserved between different *Clavibacter* species (Fig. S2). The 1,258 orthologous genes were concatenated and a maximum likelihood approach employed to generate a *Clavibacter* phylogeny (Fig. 1). *Cm* strains formed a monophyletic clade strongly supported by a high bootstrap value (100%) and was clearly separated from other *Clavibacter* species. The phylogenetic analyses also refined the genetic relationship among the *Cm* strains (Fig. 1). The relatedness among *Cm* strains does not correlate with geographical location or year of isolation. This is likely due to the global production and distribution of tomato seed, making it extremely hard to trace geographic origins of strains or sources. Collectively, these data demonstrate that *Cm* strains are not clonal and support the concept of multiple independent introductions, consistent with seed-borne disease.

All *Clavibacter* genomes contained one circular chromosome approximately 3.2 Mb in size, encoding an average of ~3,200 protein-coding sequences per genome (Table 1, Fig. S1 BC). Plasmids with similarity to pCM1 and pCM2, which are found in reference strain NCPPB382, were detected based on similarity in their plasmid backbone (replication/partitioning genes) and conjugal transfer genes. *Cm* strains possessed at least one virulence plasmid of the pCM1 or pCM2 type (Fig. 1) (Meletzus et al. 1993). Of the 34 *Cm* genomes analyzed, 33 possessed pCM1-like plasmids, while 29 possessed pCM2-like plasmids (Fig. 1). Both plasmids are required for virulence in NCPPB382 (Meletzus et al. 1993). However, we have identified multiple *Cm* strains with only pCM2 that retain pathogenicity (Thapa et al. 2017), Fig. 1). In addition, the *Cm* strain NZ2541 lacks pCM1, but is still pathogenic on tomato (Fig. 1&2). We compared the genomes of reference *Cm* strain NCPPB382 with strains NZ2541, which lacks pCM1, and CASJ001, which lacks pCM2. Genome visualizations were performed using circos (Fig. S1BC) (Krzywinski et al. 2009). Although the plasmid components of *Cm* strains NCPPB382, NZ2541 and CASJ001 differ, they display a high level of chromosomal sequence similarity (Fig. S1 BC). We also detected some instances of genome rearrangement between *Cm* strains, including CASJ001 and NZ2541 (Fig. S1BC).

Although there is significant variation within pCM1-like plasmids, the *celA* gene is conserved in the pCM1 plasmids of all *Cm* strains analyzed (Fig. 1). Previously, it was demonstrated that *celA* is required for the induction of bacterial wilt of tomato in

NCPPB382 and CASJ002 (Thapa et al. 2017; Jahr et al. 2000). However, NZ2541 also causes canker symptoms on tomato and grows to similar levels as NCPPB382 and CASJ002, even though it lacks pCM1 and *celA* (Fig. 2). Strain NZ2541 contains pCM2 in addition to a novel plasmid that we have named pCM3 (Fig. 2A). These data highlight the genetic variation present in *Cm* plasmid types.

Comparative genomics and design of *C. michiganensis* specific primers. Our genomic analyses of diverse *Clavibacter* strains indicated that several of the existing *Cm* specific primers were likely non-specific, considering the differences in plasmid content and composition (Table S2). We analyzed eight primer pairs for specificity in detecting *Cm:* PFC3/5, P5/6, ppA, tomA, celA, Cmm3/4, Cmm5/6, and PSA (Drier et al. 1995; Jahr et al., 2000; Kleitman et al. 2008; Pastrik and Rainey, 1999; Santos et al. 1997). In order to determine the specificity of commonly used published primers, we tested them on a panel of 110 bacterial strains, including 75 *Cm* strains, 20 other *Clavibacter* strains, 6 *Pseudomonads*, and 9 *Xanthomonads* (Table S1 and S2). Of the eight tested primer pairs, only the tomAF/R pair, amplifying the *tomA* tomatinase on the *chp/tomA* genomic island specifically detected all pathogenic *Cm* strains, but not other *Clavibacter* species/subspecies (Table S2).

Given the genetic diversity of *Cm* and variation in plasmid types, we sought to identify multiple genetic loci conserved among and specific to *Cm*. To identify *Cm*-specific loci, we performed whole genome comparisons of 145,929 genes from 45 *Clavibacter* genomes. We identified *Cm*-specific loci and tested ten primer pairs whose amplicons

306 could be separated by size by gel electrophoreses after amplification in a single
307 multiplex PCR (mPCR) assay. From these primer pairs, two targets were chosen: *rhuM*
308 (hypothetical protein) (CMM_2694), and *tomA* (tomatinase) (CMM_0090), were
309 ultimately selected due to their high amplification efficiency (Fig. 1, 3A). The sequences
310 of these genes were compared individually to the nr and the microbial genome
311 databases of the NCBI BLAST web site to confirm their specificity. These results
312 revealed that *tomA* and *rhuM* are highly conserved among *Cm* and similar nucleotide
313 sequences are not found in other bacteria (Fig. 3). The amplification specificity of the
314 *tomA* and *rhuM* primer pairs was confirmed in PCR tests with DNA of a bacterial panel
315 by PCR (Table S1).
316
317 The mPCR assay targets three loci, two of which are *Cm* specific, *tomA, rhuM*, and a
318 third amplifying a region of the 16S rRNA in bacteria and plant mitochondria that serves
319 as an internal positive control. Primer sets were designed with possess similar melting
320 temperatures (58- 63C), direct the amplification of the conserved *Cm* sequences, and
321 produce PCR product sizes so that the amplicons could be readily distinguished. The
322 primer sets directed the amplification of a 1000bp *rhuM* fragment, a 630bp *tomA*
323 fragment, and a 415bp 16S rRNA fragment (Fig. 3B).
324
325 **Validation of the *C. michiganensis* mPCR assay.** The specificity of the mPCR assay
326 was determined by analyzing 75 *Cm* strains, 20 other *Clavibacter* strains, and 15 non-
327 *Clavibacter* species (Table S1). The RhuMF/R and TomA-F1/R1 primer sets directed
328 the amplification of the target sequences from *Cm* strains, but not from DNA of other

Clavibacter species or other bacteria (Fig. 3B). As shown in Fig. 3B, the mPCR amplified the three expected size fragments of 1000 bp (*rhuM*), 630 bp (*tomA*), and 415 bp (16S rRNA) when purified *Cm* DNA was used as a template. No false positives were detected with *Clavibacter* endophytes, *Clavibacter* pathogens on other crops, or other bacterial pathogens of tomato, demonstrating the specificity of the mPCR assay (Fig. 3B, Fig. 4B). In contrast, the 415 bp 16S rRNA fragment was amplified from tomato and all bacterial strains tested, validating the utility of the internal positive control. The identity of selected PCR products was confirmed by sequencing.

We next evaluated the limit of detection of the mPCR assay. The detection limit was 0.01 ng with CTAB purified genomic DNA from *Cm* strain CASJ002 grown in pure culture (Fig. S3A). We also analyzed the ability to detect *Cm* after serially diluting cultures of *Cm* strain CASJ002 followed by simple DNA extraction using the lysis method and were able to reliably detect 500 cells (Fig. S3B). Similar detection levels were found for *Cm* strain NCPPB382 with both extraction methods (data not shown).

Multiplex PCR can detect *C. michiganensis* in contaminated plant and seed samples. To determine the capacity of the mPCR assay for detection of *Cm in planta*, we extracted DNA with the CTAB method from inoculated tomato plants (cultivar Early Pak 7) under greenhouse conditions. Samples were collected at seven and 14 days post-inoculation from 1, 2, 3, 6, 9, and 15 cm above the inoculation site, including from leaves. Samples were ground and split in two for DNA extraction and determining bacterial titers (Fig. 4). *Cm* was detected from all stem and leaf tissue samples of

inoculated plants at both time points, whereas it was not detected from mock-inoculated controls (Fig. 4). Under our experimental conditions, wilting disease symptoms were visible in five to seven days post-inoculation. For infected samples, the expected amplicons of 1000 bp (*rhuM*), 630 bp (*tomA*), and 415 bp (16S rRNA) were observed, whereas only the 415 bp fragment was observed for DNA extracted from mock-inoculated plants (Fig. 4BC i). We did not collect samples at 15 cm above the inoculation site at 7 days post-inoculation as the plants were too small in height at this time points. However, the mPCR assay was able to detect *Cm* up to 9 cm above the inoculation site at 7 days post-inoculation (Fig. 4B i). The mPCR assay was able to detect *Cm* up to 15 cm above the site of infection at 14 days post-inoculation, which corresponded to 8 log colony forming units per milligram plant tissue (Fig. 4C ii).

The limit of detection of the mPCR assay in plant samples was also determined. DNA was isolated from tomato plants with the CTAB method (cultivar Early Pak 7). The bacterial titers were experimentally determined as previously described and used to calculate the limit of detection (Fig. 4D). The level of sensitivity for *in planta* detection was similar to that obtained with the bacterial cultures, and we were able to detect *Cm* in 0.02 ng of plant DNA, corresponding to ~10^3 *Cm* cells per PCR reaction (Fig. 4D). The sensitivity of the mPCR was also evaluated by spiking *Cm* CASJ002 DNA into uninfected tomato DNA samples in the ratios (bacterial: tomato DNA) 1:1, 1:5, 1:25, 1:125, 1:250, 1:312.5 and with 100 ng/ul dilution in each PCR reaction. In these experiments, *Cm* was detected down to 1:250 of *Cm*:tomato DNA dilution (Fig. 4E).

We also tested the mPCR assay for detection of *Cm* in DNA extracted from artificially infested seeds. Different concentrations of *Cm* strain CASJ002 were used to infest tomato seeds, followed by DNA extraction and mPCR (Fig. 5A). The mPCR assay was able to detect *Cm* with bacterial contamination in DNA isolated from a sample of 10 infested seed (Fig. 5A). Next, we evaluated the mPCR for detecting *Cm* contamination after spiking contaminated seed into clean seed lots. Seeds infested with $5x10^7$ *Cm* CASJ002 were mixed with 10,0000 clean seeds in different ratios followed by DNA extraction and mPCR. The mPCR was able to detect 0.3% incidence of *Cm* contamination in seeds (Fig. 5B).

Evaluation of multiplex PCR detection from recent *Cm* field strains. In 2016, there were bacterial canker occurrences in in the California Central Valley (Table S3). Tomato samples exhibiting symptoms of wilt, canker, and bird's eye spots on fruit were collected from fields and assayed for the presence of *Cm* (Table S3; Fig. S4). The three fragments *rhuM* (1000 bp), *tomA* (630 bp) and 16S rRNA (415 bp) were amplified from samples with canker symptoms, whereas only the fragment corresponding to 16S rRNA was amplified from healthy plants (Fig. 6A). With the mPCR assay, *Cm* was detected in all 21 samples with canker symptoms tested, consistent with validation of *Cm* from the samples on D2 medium. Furthermore, we verified the pathogenicity of selected *Cm*-like colonies by inoculating tomato Early Pak 7. Samples that tested positive using the mPCR assay generated *Cm*-like colonies on D2 media. These colonies also caused bacterial canker disease symptoms and grew to similar levels as previously characterized strains, including CASJ002 (Fig. 6BC). Taken together, these data

demonstrate the specificity and utility of mPCR detection of *Cm* in field samples and diagnosis of bacterial canker.

DISCUSSION

In this study, we used a comparative genomics approach to analyze strains and species of the genus *Clavibacter* and used this information to develop a specific detection assay for *Cm*, the causal agent of bacterial canker of tomato. Here, we have identified orthologous loci that are distributed among *Clavibacter* species. A total of 1,258 core orthologous loci were used to generate high resolution *Clavibacter* phylogeny, which clearly demonstrates that *Cm* is separate from other *Clavibacter* species. The highly similar orthologous loci among the *Cm* also indicates their conserved function. The analyses of the *Cm* genomes supports the monophyletic nature of this pathogen. However, *Cm* strains exhibit genetic diversity when compared with each other. The chp/tomA pathogenicity island is conserved among all the *Cm* strains and has been previously demonstrated to be a critical virulence factor in *Cm* strain NCPPB382 (Kaup et al., 2005; Gartemann et al., 2008). Significant differences in plasmid content and composition were observed, with is consistent with previous findings of plasmid diversity in *Cm* strains from California and New York (Thapa et al., 2017; . Our previous study on California *Cm* strains also had similar findings (Thapa et al., 2017). Differences in genomic island, secreted protein and CAZymes were also observed (Tancos et al., 2015; Thapa et al., 2017).

The *Cm* strains in our present phylogeny do not cluster based on geographical origin or the year of isolation, which is consistent with seed-borne disease and separate

introduction events. A similar distribution pattern was observed for *Cm* outbreaks in Turkey (Baysal et al. 2011). This is also supported by studies reporting that a single source of introduction by seed which persisted afterwards in the region (de León et al. 2009; Jacques et al. 2012; Kawaguchi et al. 2010). Our comparative genomics approach further identified loci that are conserved in all *Cm* and absent in other *Clavibacter* species and bacterial pathogens of tomato. Of 10 loci examined, two were used to design a *Cm* detection assay based on mPCR. We demonstrate the power of a genomics approach for comparing a large number of related bacterial genomes to identify specific genetic markers for pathogen detection.

The best way to manage bacterial canker disease of tomato is using a seed source free of *Cm* contaminants, extensive sanitation of equipment, removal of crop debris, and crop rotation in the field. Effective detection methods are critical for early detection of outbreaks and for detection of contaminated seed lots. Even low levels of *Cm* contamination (0.01%) from seed to seedling can cause a disease epidemic under favorable conditions (Chang et al. 1991). Therefore, accurate detection of *Cm* in seed is an important component of disease management, especially given the prevalence of diverse endophytic *Clavibacter* strains associated with tomato. The most common detection assays for *Cm* are serological or PCR-based, following isolation of suspect bacteria by growth on semi-selective media (Dreier et al. 1995; Louws et al. 1998; Kaneshiro et al. 2006; Milijasevic-Marcic et al. 2012; Alvarez, 2005; ISHI, 2017). The serological techniques are not reliable due to their cross reaction with other *Clavibacter* species (Franken et al. 1993; Dreier et al. 1995; Jacques et al. 2012). Most of the

present PCR assays for *Cm* are based on plasmid targets (Alvarez, 2005; Kleitman et al. 2008; Dreier et al. 1995; Santos et al. 1997). Whereas all pathogenic *Cm* strains possess at least one virulence plasmid, we have identified significant variation in *Cm* plasmid content and profiles (Thapa et al. 2017). For example, the NZ2541 strain lacks pCM1 but is able to cause disease on tomato. Furthermore, only 85% of the *Cm* strains sequenced possess pCM2-like plasmids. Other studies have reported *Cm* plasmid variation in field strains and plasmid loss under laboratory conditions (Kleitman et al. 2008; Thapa et al. 2017; Tancos et al. 2015). For example, the frequently used CM5/CM6 primer pairs which targets the pat-1 locus in pCM2 only directed the amplification of the target fragment from 75% of tested *Cm* strains, while the CM3/CM4 primer pair which targets the ppaJ locus in pCM1 directed the amplification of the target fragment from 83% of tested *Cm* strains (Louws et al. 1999; Hadas et al. 2005; Santos et al., 1999). In the present study, these primers showed both false positive and negative results (Table S2). Therefore, the published primers targeting plasmids for detection of *Cm* are not sufficiently reliable.

Diagnostic tests for pathogen presence should possess specificity, sensitivity, and reliability. In our mPCR diagnostic assay for *Cm*, we included three primer pairs that target two *Cm* chromosomal loci (*tomA* and *rhuM*), as well as the 16S rRNA gene as an internal control that amplifies bacterial or plant mitochondrial DNA. *TomA* is found in the *chp/tomA* pathogenicity island, which is present in all pathogenic *Cm* strains, but is absent in *Clavibacter* endophytes and *Clavibacter* species infecting other crops (Kaup et al. 2005; Thapa et al. 2017). *TomA* encodes a putative tomatinase, which

deglycosylates the antibacterial saponin α-tomatine (Kaup et al. 2005). The *rhuM* gene is also conserved in all *Cm* strains (Table S1). RhuM is predicted to be involved in virulence in other bacterial pathogens based on its presence on the PAI-3 of *Salmonella* spp., but its function in *Cm* virulence remains to be determined (Bertelloni et al. 2017). We verified the specificity of the mPCR assay against 75 *Cm* strains, 20 other *Clavibacter* strains and 15 other bacterial plant pathogens. We detected as little as ~10^3 CFU/ml of *Cm,* which is a similar level of sensitivity also reported for PCR in other studies (Dreier et al. 1995; Luo et al. 2008). We were also able to detect *Cm* from seed after artificial contamination. The mPCR assay reported herein proved to be a sensitive, accurate, and reliable diagnostic technique for *Cm* in infected plant materials.

Recent advancements in genome sequencing and high throughput computing have enabled pathogen tracking and development of pathogen detection assays (Islam et al. 2016; Bueno-Sancho et al. 2017; Langlois et al. 2017; Savory et al. 2017). For instance, comparative analysis has been used to identify emergence of *Magnaporthe oryzae* lineage and its origin in Bangladesh, and track and define *Puccinia striiformis* population structure and dispersal on a global scale (Islam et al. 2016; Bueno-Sancho et al. 2017). It has also been used for real time monitoring of *Phytophthora infestans* population for virulence and application of preventive fungicide (Li et al. 2012). Comparative genomics has also enabled specific detection and identification of different plant pathogenic bacteria (Langlois et al. 2017; Savory et al. 2017). Our results demonstrate the successful application of comparative genomics to identify specific loci for accurate detection of *Cm*. The genetic markers obtained from whole genome comparative

analysis used in the mPCR assay described here are specific, sensitive, and reliable for the detection of the *Cm,* both in infected tomato plants and seeds. The adoption of this assay by the seed industry and diagnostic laboratories will facilitate rapid screening of infected plants and seeds for routine diagnostics.

Acknowledgements

We thank CIRM-CFBP (https://www6.inra.fr/cirm_eng/**)** for providing part of *Cm* strains used in this study.

LITERATURE CITED

Alvarez, A. M., Kaneshiro, W. S., and Vine, B. G. 2005. Diversity of *Clavibacter michiganensis* subsp. *michiganensis* populations in tomato seed: What is the significance? Acta Hortic. 695:205-213.

Anonymous 2000. European Union Council directive 2000/29/EC on protective measures against the introduction into the Community of organisms harmful to plants or plant products and against their spread within the Community. Off. J. Eur. Commun. L. 169:33.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., Pevzner, P. A. 2012. SPAdes: a new

genome assembly algorithm and its applications to single-cell sequencing. J. comput. biol. 19:455-477.

Baysal, O., Mercati, F., Ikten, H., Yildiz, R. C., Carimi, F., Aysan, Y., Jaime, A., da Silva, T. 2011, *Clavibacter michiganensis* subsp *michiganesis*: Tracking strains using their genetic differentiations by ISSR markers in Southern Turkey. Physiol. Mol. Plant Pathol. 75:113-119.

Bertelloni, F., Tosi, G., Massi, P., Fiorentini, L., Parigi, M., Cerri, D., Ebani V. V. 2017. Some pathogenic characters of paratyphoid *Salmonella enterica* strains isolated from poultry. Asian Pac. J. Trop. Med. 10:1161-1166.

Bertolini, E., Olmos, A., Lopez, M. M. and Cambra, M. 2003. Multiplex Nested Reverse Transcription-Polymerase Chain Reaction in a Single Tube for Sensitive and Simultaneous Detection of Four RNA Viruses and *Pseudomonas savastanoi* pv. *savastanoi* in Olive Trees. Phytopathology 93:286-292.

Bueno-Sancho, V., Persoons, A., Hubbard, A., Cabrera-Quio, L. E., Lewis, C. M., Corredor-Moreno, P., Bunting, D. C. E., Ali, S., Chng, S., Hodson, D. P., Madariaga Burrows, R., Bryson, R., Thomas, J., Holdgate, S., Saunders, D. G. O. 2017. Pathogenomic Analysis of Wheat Yellow Rust Lineages Detects Seasonal Variation and Host Specificity. Genome Biol. Evol. 9:3282-3296.

Chang, R. J., Ries, S. M. and Pataky, J. K. 1991 Dissemination of *Clavibacter michiganensis* subsp *michiganensis* by Practices Used to Produce Tomato Transplants. Phytopathology 81:1276-1281.

de León, L., Rodriguez, A., Llop, P., Lopez, M. M. and Siverio, F. 2009. Comparative study of genetic diversity of *Clavibacter michiganensis* subsp *michiganensis*

isolates from the Canary Islands by RAPD-PCR, BOX-PCR and AFLP. Plant Pathol 58:862-871.

de León, L., Siverio, F., López, M. M. and Rodríguez, A. 2011. *Clavibacter michiganesis* subsp. *michiganensis*, a Seedborne Tomato Pathogen: Healthy Seeds Are Still the Goal. Plant Disease 95:1328-1338.

Dreier, J., Bermpohl, A. and Eichenlaub, R. 1995. Southern Hybridization and Pcr for Specific Detection of Phytopathogenic *Clavibacter michiganensis* subsp m*ichiganensis*. Phytopathology 85:462-468.

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., Gartemann, K.-H. and Burger, A. 2006. *Clavibacter michiganensis*, a group of gram-positive phytopathogenic bacteria. In: *Plant-Associated Bacteria.* (Gnanamanickam, S. S., ed.). Dordrecht: Springer Netherlands, pp. 385-421.

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

Fatmi, M. and Schaad, N. W. 2002. Survival of *Clavibacter michiganensis* ssp *michiganensis* in infected tomato stems under natural field conditions in California, Ohio and Morocco. Plant Pathol. 51:149-154.

Fatmi, M., Bolkan, H., and Schaad, N. W. 2017. Detection of Clavibacter michiganense subsp. michiganensis in tomato seed. Pages 111-117 In: Detection of Plant-Pathogenic Bacteria in Seed and other Planting Material. 2nd edition. M. Fatmi,

559 R. R. Walcott, and N. W. Schaad, eds. The American Phytopathological Society.
560 St. Paul, MN. 360 pp.
561 Franken, A. A. J. M., Kamminga, G. C., Snijders, W., Vanderzouwen, P. S. and
562 Birnbaum, Y. E. 1993. Detection of *Clavibacter michiganensis* ssp. *mchiganensis*
563 in Tomato Seeds by Immunofluorescence Microscopy and Dilution Plating. Neth.
564 J. Plant Pathol. 99:25-137.
565 Gartemann, K.H., Abt, B., Bekel, T., Burger, A., Engemann, J., Flügel, M., Gaigalat, L.,
566 Goesmann, A., Gräfen, I., Kalinowski, J. and Kaup, O., 2008. The genome
567 sequence of the tomato-pathogenic actinomycete Clavibacter michiganensis
568 subsp. michiganensis NCPPB382 reveals a large island involved in
569 pathogenicity. Journal of Bacteriology 190:2138-2149.
570 Hadas, R., Kritzman, G., Klietman, F., Gefen, T. and Manulis, S. 2005. Comparison of
571 extraction procedures and determination of the detection threshold for
572 *Clavibacter michiganensis* ssp. *michiganensis* in tomato seeds. Plant Pathol.
573 54:643-649.
574
575 ISHI. 2017. Method for the Detection of *Clavibacter michiganensis* subsp.
576 *michiganensis* on Tomato seed (Version 4.3.1). International Seed Health
577 Initiative. https://www.worldseed.org/wp-
578 content/uploads/2017/07/Tomato_Cmm_July2017.pdf
579 Islam, M. T., Croll, D., Gladieux, P., Soanes, D. M., Persoons, A., Bhattacharjee, P.,
580 Hossain, M. S., Gupta, D. R., Rahman, M. M., Mahboob, M. G., Cook, N., Salam,
581 M. U., Surovy, M. Z., Sancho, V. B., Maciel, J. L., NhaniJúnior, A., Castroagudín,

582 V. L., Reges, J. T., Ceresini, P. C., Ravel. S., Kellner, R., Fournier, E., Tharreau,
583 D., Lebrun, MH., McDonald, B. A., Stitt, T., Swan, D., Talbot, N. J., Saunders, D.
584 G., Win, J., Kamoun, S. 2016. Emergence of wheat blast in Bangladesh was
585 caused by a South American lineage of *Magnaporthe oryzae*. Bmc Biol 14:84.
586 Jacques, M. A., Durand, K., Orgeur, G., Balidas, S., Fricot, C., Bonneau, S., Quillévéré,
587 A., Audusseau, C., Olivier, V., Grimault, V., Mathis R. 2012. Phylogenetic
588 analysis and polyphasic characterization of *Clavibacter michiganensis* strains
589 isolated from tomato seeds reveal that nonpathogenic strains are distinct from *C.*
590 *michiganensis* subsp. *michiganensis*. Appl. Environ. Microbiol. 78:8388-8402.
591 Jahr, H., Dreier, J., Meletzus, D., Bahro, R. and Eichenlaub, R. 2000. The endo-beta-
592 1,4-glucanase CelA of *Clavibacter michiganensis* subsp. *michiganensis* is a
593 pathogenicity determinant required for induction of bacterial wilt of tomato. Mol.
594 Plant-microbe Interact. 13:703-714.
595 Kado, C. I. and Heskett, M. G. 1970 Selective Media for Isolation of *Agrobacterium*,
596 *Corynebacterium*, *Erwinia*, *Pseudomonas*, and *Xanthomonas*. Phytopathology
597 60:969-976.
598 Kaneshiro, W. S., Mizumoto, C. Y. and Alvarez, A. M. 2006. Differentiation of
599 *Clavibacter michiganensis* subsp *michiganensis* from seed-borne saprophytes
600 using ELISA, Biolog and 16S rDNA sequencing. E. J. Plant Pathol 116:45-56.
601 Kaup, O., Grafen, I., Zellermann, E. M., Eichenlaub, R. and Gartemann, K. H. 2005.
602 Identification of a tomatinase in the tomato-pathogenic actinomycete *Clavibacter*
603 *michiganensis* subsp. *michiganensis* NCPPB382. Mol. Plant-microbe Interact.
604 18:1090-1098.

605 Kawaguchi, A., Tanina, K. and Inoue, K. 2010. Molecular typing and spread of
606 *Clavibacter michiganensis* subsp *michiganensis* in greenhouses in Japan. Plant
607 Pathol. 59:76-83.
608 Kirchner, O., Gartemann, K. H., Zellermann, E. M., Eichenlaub, R. and Burger, A. 2001.
609 A highly efficient transposon mutagenesis system for the tomato pathogen
610 *Clavibacter michiganensis* subsp. *michiganensis*. Mol. Plant-microbe Interact.
611 14:1312-1318.
612 Kleitman, F., Barash, I., Burger, A., Iraki, N., Falah, Y., Sessa, G., Weinthal, D.,
613 Chalupowicz, L., Gartemann, K. H., Eichenlaub, R., Manulis-Sasson, S., 2008.
614 Characterization of a *Clavibacter michiganensis* subsp *michiganensis* population
615 in Israel. E. J. Plant Pathol. 121:463-475.
616 Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.
617 J., Marra, M. A. 2009. Circos: an information aesthetic for comparative genomics.
618 Genome research 19:1639-1645.
619 Kuck, P. and Meusemann, K. 2010. FASconCAT: Convenient handling of data matrices.
620 Mol. Phylogent. Evol. 56:1115-1118.
621 Langlois, P. A., Snelling, J., Hamilton, J. P., Bragard, C., Koebnik, R., Verdier, V.,
622 Triplett, L. R., Blom, J., Tisserat, N, A., Leach, J. E. 2017. Characterization of the
623 *Xanthomonas translucens* Complex Using Draft Genomes, Comparative
624 Genomics, Phylogenetic Analysis, and Diagnostic LAMP Assays. Phytopathology
625 107:519-527.
626 Li, L., Stoeckert, C. J., Jr. and Roos, D. S. 2003. OrthoMCL: identification of ortholog
627 groups for eukaryotic genomes. Genome research 13:2178-2189.

628 Li, W. and Godzik, A. 2006. Cd-hit: a fast program for clustering and comparing large
629 sets of protein or nucleotide sequences. Bioinformatics 22:1658-1659.
630 Li, Y., van der Lee, T. A. J., Evenhuis, A., van den Bosch, G. B. M., van Bekkum, P. J.,
631 Forch, M. G., van Gent-Pelzer, M. P. E., van Raaij, H. M. G, Jacobsen, E.,
632 Huang S. W., Govers, F., Vleeshouwers V.G. A. A., Kessel, G. J.2012.
633 Population Dynamics of Phytophthora infestans in the Netherlands Reveals
634 Expansion and Spread of Dominant Clonal Lineages and Virulence in Sexual
635 Offspring. G3-Genes Genom Genet 2:1529-1540.
636 Li, X., Tambong, J., Yuan, K. X., Chen, W., Xu, H., Levesque, C. A., DeBoer, S. H.
637 2018. Re-classification of Clavibacter michiganensis subspecies on the basis of
638 whole-genome and multi-locus sequence analyses. Int. J. Syst. Evol. Microbiol.
639 68:234-240.
640 Louws, F., Rademaker, J. and de Bruijn, F. 1999. The three DS of PCR-based genomic
641 analysis of phytobacteria: Diversity, Detection, and Disease Diagnosis. Ann.
642 review phytopathol. 37:81-125.
643 Louws, F. J., Bell, J., Medina-Mora, C. M., Smart, C. D., Opgenorth, D., Ishimaru, C. A.,
644 Hausbeck, M. K., de Bruijn, F. J., Fulbright, D. W. 1998. rep-PCR-mediated
645 genomic fingerprinting: A rapid and effective method to identify Clavibacter
646 michiganensis. Phytopathology 88:862-868.
647 Loytynoja, A. and Goldman, N. 2005. An algorithm for progressive multiple alignment of
648 sequences with insertions. PNAS 102:10557-10562.

649 Luo, L. X., Walters, C., Bolkan, H., Liu, X. L. and Li, J. Q. 2008. Quantification of viable
650 cells of Clavibacter michiganensis subsp michiganensis using a DNA binding dye
651 and a real-time PCR assay. Plant Pathol. 57:332-337.
652 Maddox, D. (1997) Regulatory needs for standarized seed health tests. 79a, pp. 81–92
653 Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., Dow, M.,
654 Verdier, V., Beer, S. V., Machado, M. A., Toth, I., Salmond, G., Foster, G. D.
655 2012. Top 10 plant pathogenic bacteria in molecular plant pathology. Mol. Plant.
656 Pathol. 13:614-629.
657 McNally, R. R., Ishimaru, C. A. and Malvick, D. K. 2016. PCR-Mediated Detection and
658 Quantification of the Goss's Wilt Pathogen *Clavibacter michiganensis* subsp.
659 *nebraskensis* Via a Novel Gene Target. Phytopathology 106:1465-1472.
660 Meletzus, D., Bermphol, A., Dreier, J. and Eichenlaub, R. 1993. Evidence for plasmid-
661 encoded virulence factors in the phytopathogenic bacterium *Clavibacter*
662 *michiganensis* subsp. *michiganensis* NCPPB382. J. Bacteriol. 175:2131-2136.
663 Milijašević-Marčić, S., Gartemann, K.-H., Frohwitter, J., Eichenlaub, R., Todorović, B.,
664 Rekanović, E., Potočnik, I. 2012. Characterization of *Clavibacter michiganensis*
665 subsp. *michiganensis* strains from recent outbreaks of bacterial wilt and canker in
666 Serbia. E. J. Plant Pathol. 134:697-711.
667 Murray, M. G. and Thompson, W. F. 1980. Rapid isolation of high molecular weight
668 plant DNA. Nucleic Acids Res. 8:4321-4325.
669 Ozdemir, Z. 2009. Development of a multiplex PCR assay for the simultaneous
670 detection of *Clavibacter michiganensis* subsp. *michiganensis*, *Pseudomonas*

syringae pv. *tomato* and *Xanthomonas axonopodis* pv. *vesicatoria* using pure clutures. J. Plant. Pathol. 91:495-497.

Rambaut, A. (2012) FigTree version 1.4. 0. *http://tree.bio.ed.ac.uk/software/figtree*.

Santos, M. S., Cruz, L., Norskov, P. and Rasmussen, O. F. 1997. A rapid and sensitive detection of *Clavibacter michiganensis* subsp. *michiganensis* in tomato seeds by polymerase chain reaction. Seed Sci. Technol. 25:581-584.

Savory, E. A., Fuller, S. L., Weisberg, A. J., Thomas, W. J., Gordon, M. I., Stevens, D. M., Creason, A. L., Belcher M. S., Serdani, M., Wiseman, M. S., Grunwald N. J., Putnam, M. L., Chang J. H. 2017. Evolutionary transitions between beneficial and phytopathogenic *Rhodococcus* challenge disease management. Elife, 6. e30925.

Schaad, N. W. and Frederick, R. D. 2002. Real-time PCR and its application for rapid plant disease diagnostics. Can. J. Plant. Pathol. 24:250-258.

Seemann, T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068-2069.

Sen, Y., van der Wolf, J., Visser, R. G. F. and van Heusden, S. 2014. Bacterial Canker of Tomato: Current Knowledge of Detection, Management, Resistance, and Interactions. Plant Disease 99:4-13.

Smith, E. F. 1910. A new tomato disease of economic importance. Science 31.

Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688-2690.

692 Tancos, M. A., Lange, H. W. and Smart, C. D. 2015. Characterizing the genetic diversity
693 of the *Clavibacter michiganensis* subsp. *michiganensis* population in New York.
694 Phytopathology 105:169-179.
695 Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E. P., Zaslavsky,
696 L., Lomsadze, A., Pruitt, K. D., Borodovsky, M., Ostell, J. 2016. NCBI prokaryotic
697 genome annotation pipeline. Nucleic Acids Res. 44:6614-6624.
698 Thapa, S. P., Han, A. R., Cho, J. M. and Hur, J. H. 2013. Multiplex PCR and DNA array
699 for the detection of *Bacillus cereus*, *Staphylococcus aureus*, *Listeria*
700 *monocytogenes*, *Escherichia coli* O157:H7, and *Salmonella* spp. targeting
701 virulence-related genes. Ann. Microbiol. 63:725-731.
702 Thapa, S. P., Miyao, E. M., Davis, R. M. and Coaker, G. 2015. Identification of QTLs
703 controlling resistance to *Pseudomonas syringae* pv. *tomato* race 1 strains from
704 the wild tomato, Solanum habrochaites LA1777. Theor. Appl. Genet. 128:681-
705 692.
706 Thapa, S. P., Park, D. H., Wilson, C., Hur, J. H. and Lim, C. K. 2012. Multiplex PCR
707 assay for the detection of *Pantoea stewartii* subsp. *stewartii* using species-
708 specific genetic markers. Australas Plant Path. 41:559-564.
709 Thapa, S. P., Pattathil, S., Hahn, M. G., Jacques, M. A., Gilbertson, R. L. and Coaker,
710 G. 2017. Genomic Analysis of *Clavibacter michiganensis* Reveals Insight Into
711 Virulence Strategies and Genetic Diversity of a Gram-Positive Bacterial
712 Pathogen. Mol. Plant-microbe Interact. 30:786-802.
713 Vincelli, P. and Tisserat, N. 2008. Nucleic acid-based pathogen detection in applied
714 plant pathology. Plant Disease 92:660-669.

715 Yasuhara-Bell, J. and Alvarez, A. M. 2015. Seed-associated subspecies of the genus
716 *Clavibacter* are clearly distinguishable from *Clavibacter michiganensis* subsp.
717 *michiganensis*. Int. J. Syst. Evol. Microbiol. 65:811-826.
718 Zaluga, J., Van Vaerenbergh, J., Stragier, P., Maes, M. and De Vos, P. 2013. Genetic
719 diversity of non-pathogenic *Clavibacter* strains isolated from tomato seeds. Syst.
720 Appl. Microbiol. 36:426-435.

721

722

723

724 **TABLE AND FIGURE LEGENDS:**

725

726 **Table 1.** Genome statistics and GenBank accession numbers of *Clavibacter* genomes

727 analyzed in this study.

species	Strains	Source	# contigs	Note	Genome size (bp)	Accession number
C. michiganensis	ARZ28	USA	434	This study	3281597	QLNE00000000
C. michiganensis	ATCC10202	-	334	This study	3275372	QLMX00000000
C. michiganensis	ATCC14456	Italy	41	This study	3296963	QLMU00000000
C. michiganensis	CFBP1465	France	66	This study	3328034	QLMV00000000
C. michiganensis	CFBP1940	Spain	35	This study	3296592	QLMW00000000
C. michiganensis	CFBP2494	Algeria	39	This study	3316980	QLMY00000000
C. michiganensis	CFBP2500	Algeria	332	This study	3305560	QLMZ00000000
C. michiganensis	CFBP5842	Brazil	127	This study	3343022	QLNA00000000
C. michiganensis	CFBP6885	France	25	This study	3327889	QLML00000000
C. michiganensis	CFBP7158	New Zealand	23	This study	3285302	QLNB00000000
C. michiganensis	CFBP7311	Morocco	46	This study	3301380	QLNC00000000
C. michiganensis	CFBP7312	China	43	This study	3319896	QLND00000000
C. michiganensis	CFBP7314	USA	30	This study	3268866	QLMM00000000
C. michiganensis	CFBP7315	USA	31	This study	3311336	QLMN00000000
C. michiganensis	CFBP7316	USA	29	This study	3258979	QLMO00000000
C. michiganensis	CFBP7488	France	55	This study	3299275	QLMP00000000
C. michiganensis	CFBP7568	USA	61	This study	3347875	QLMQ00000000
C. michiganensis	CFBP7589	Belgium	35	This study	3316741	QLMR00000000
C. michiganensis	NZ1811	New Zealand	333	This study	3273907	QLMS00000000
C. michiganensis	NZ5026	USA	263	This study	3283188	QLMT00000000
C. michiganensis	NZ2541	UK	292	This study	3321838	QODA00000000
C. michiganensis	CASJ001	USA	4	GenBank	3423339	MDHB00000000
C. michiganensis	CASJ002	USA	3	GenBank	3423641	MDHC00000000
C. michiganensis	CASJ003	USA	315	GenBank	3294503	MDHD00000000
C. michiganensis	CASJ004	USA	477	GenBank	3351833	MDHE00000000
C. michiganensis	CASJ005	USA	321	GenBank	3284480	MDHF00000000
C. michiganensis	CASJ006	USA	4	GenBank	3499162	MDHG00000000
C. michiganensis	CASJ007	USA	11	GenBank	3400049	MDHH00000000
C. michiganensis	CASJ008	USA	39	GenBank	3392817	MDHI00000000
C. michiganensis	LMG7333^T	Hungary	5	GenBank	3391512	NZ_MZMP00000000.1

C. michiganensis	VKMAc	Hungary	13	GenBank	3295160	NZ_FVZG00000000.1
C. michiganensis	CAYO001	USA	3	GenBank	3481307	MDHL00000000
C. michiganensis	CA00001	USA	120	GenBank	3625805	MDHK00000000
C. michiganensis	CA00002	USA	4	GenBank	3371744	MDHM00000000
C. michiganensis	NCPPB382	UK	-	GenBank	339524	AM711867.1
C. sepedonicus	ATCC33113	-	1	GenBank	3258645	NC_010407.1
C. insidiosus	LMG3663	USA	3	GenBank	3387165	NZ_MZMO00000000.1
C. capcisi	PF008	South Korea	3	GenBank	3056296	NZ_CP012573
C. tessellarius	ATCC33566	USA	2	GenBank	3318535	MZMQ00000000.1
C. nebraskensis	NCPPB2581	USA	1	GenBank	3063596	NC_020891.1
Clavibacter spp.	CFBP7494	Chile	15	GenBank	3425817	MDJW00000000
Clavibacter spp.	CFBP7576	-	372	GenBank	3399000	MDJX00000000
Clavibacter spp.	CFBP8017	Netherlands	65	GenBank	3284576	MDJY00000000
Clavibacter spp.	CFBP8019	Netherlands	18	GenBank	3076288	MDJZ00000000
Clavibacter spp.	LMG26808	Netherlands	70	GenBank	3420753	NZ_AZQZ01000000

728

729 ATCC- American Type Culture Collection, CFBP = French Collection of Plant-
730 associated Bacteria, LMG - Laboratorium voor Microbiologie, ^{T} - Type strain.

Figure 1. Genome assisted phylogeny of the genus *Clavibacter*. A total of 1,258 orthologous genes from each strain were concatenated. A maximum likelihood approach was used to generate the phylogeny with 1000 bootstrap replicates. Bootstrap values are indicated at each node. Columns indicate the presence of the pCM1 plasmid, the pCM2 plasmid, the *chp/tomA* genomic island, and *RhuM*. Phenotype = blue indicates pathogenicity on tomato, Host = plants bacteria were isolated from if known, E = endophyte of tomato.

Figure 2. *C. michiganensis* strain NZ2541 lacks the pCM1 plasmid and *celA* but is pathogenic on tomato. A) Comparison of NCPPB382's pCM1 plasmid with the new plasmid found in NZ2541, pCM3. White = hypothetical proteins, red = secreted proteins, yellow = transcriptional regulator, violet = conjugal transfer protein, pink = *celA*. **B)** Representative bacterial wilt and canker symptoms induced by *C. michiganensis* strain in tomato 7 days post-inoculation (dpi). **C)** Bacterial titers in infected tomato stems 7 dpi. Three-week-old tomato plants (*Solanum lycopersicum* cv. Early Pak 7) were inoculated by stabbing tomato stems with a needle and delivering 5 µl of at 10^7 CFU/ml of bacteria. Bacterial titers were determined by grinding sterilized stem segments 7 dpi and plating on D2 semi-selective media. Statistical differences were detected by a two-tailed t-test compared and; asterisks indicate P = 0.001.

Figure 3. *C. michiganensis* primer design and multiplex PCR. A) The genetic architecture of flanking genes surrounding the *tomA* and *rhuM* loci. The *chp/tomA* cluster, including *tomA,* is only present in pathogenic *C. michiganensis.* RhuM is only

present in *C. michiganensis* tomato pathogens. Primers used to amplify each gene are shown in grey. **B)** Agarose gel electrophoresis of multiplex PCR assay products. Total genomic DNA was used as a template with the RhuM-F/R primers targeting the *rhuM* gene, the TomA-F1/R1 primers targeting the *tomA* gene, and the 16sR-F/R primers targeting 16S rRNA in either bacteria or plants. Lanes 1-5, endophytic *Clavibacter* strains (CASJ009, CFBP8017, CFBP8019, CFBP7576, and CFBP7494); lanes 6-39, pathogenic *C. michiganensis* (CFBP7311, CFBP1940, CFBP5842, CFBP7568, CFBP7488, CFBP2500, CFBP7158, CFBP7312, CFBP7316, CFBP7314, CFBP6885, CFBP7315, CFBP7589, NCPPB382, ATCC14456, ATCC10206, NZ2541, NZ1811, NZ5026, CASJ001, CASJ002, CASJ003, CASJ004, CASJ005, CASJ008, CA00001, CAYO001, AZ-28, NZ2550, 285D, 12B, 13C1, Cm4, and Cm5); lanes 40-43, *C. sepedonicus* (NZ2535, SB109, INM-1, and BRR7); lane 44, *C. insidiosus* (CIC266); lanes 45-46, *C. tessellarius* (CIC021,CIC022); lanes 47-48, *Xanthomonas euvesicatoria* (RL677, HB1); lanes 49-51 *Pseudomonas syringae* pv. *tomato* (PstA9, Pst18, and Pst838-8); lanes 52-54, *E. coli*, *Candidatus* Liberibacter asiaticus HHCA, and uninfected Tomato DNA. Lane M = 1kb DNA ladder.

Figure 4. Detection of *C. michiganensis* in infected plant materials. A). Three-week-old tomato plants (*Solanum lycopersicum* cv. Early Pak 7) were stabbed with a needle and inoculated with 5 µl of 10^7 CFU/ml suspension of CASJ002. The diagram indicates points of sample collection above the *C. michiganensis* inoculation site. **B-i, ii)** Detection of *C. michiganensis* by multiplex PCR and bacterial titers in inoculated tomato plants 7 days post-inoculation (dpi). DNA was extracted from 100 mg of plant tissue with

the CTAB method, re-suspended in 50 µl of water, and 2 µl of DNA was used per multiplex PCR reaction. **C-i, ii)** Detection of *C. michiganensis* by multiplex PCR and bacterial titers in CASJ002 inoculated tomato plants at 14 dpi. Bacterial titers, DNA extraction and multiplex PCR were performed as described in B. **D)** The detection limit for CASJ002 after CTAB DNA extraction from tomato stems. DNA was extracted from plant samples 9 cm above the site of inoculation at 7 dpi. DNA samples were diluted and the total DNA concentration in each multiplex PCR reaction is indicated above each lane. **E)** Detection of *C. michiganensis* by multiplex PCR in serial dilutions of CASJ002 DNA (50ng) and healthy tomato plant DNA (50 ng) mixed in different ratios using 100 ng/ul in a 25 µl PCR reaction. Numbers indicate bacterial:plant DNA ratios. Lane M, size marker (1kb DNA ladder).

Figure 5. Detection of *C. michiganensis* in seeds. A) *C. michiganensis* multiplex PCR detection from contaminated tomato seed. Tomato seeds were contaminated with CASJ002 at the indicated concentrations. DNA was directly extracted from 10 seeds, the pellet re-supended in 50 ul, and 2 ul of purified DNA was used in each multiplex PCR reaction. **B)** Detection of *C. michiganensis* by multiplex PCR after spiking contaminated seed with clean seed lots consisting of 10,000 seed. Seeds infested with $5x10^7$ CASJ002 were mixed with clean seeds in different ratios followed by DNA extraction and multiplex PCR. Seeds were incubated with 120 ml of PBS, crushed in a stomacher, and 5-10 ml was used for DNA extraction. The pellet was re-suspended in 100-200 ul and 2-4 ul of purified DNA was used in each PCR reaction. Infested: clean seed ratios are indicated above each lane. Lane1, 1kb DNA ladder.

800

801 **Figure 6. Multiplex PCR confirmed the presence of *C. michiganensis* in samples**
802 **collected from commercial production fields. A)** Detection of *C. michiganensis* by
803 multiplex from field samples. Purified Clavibacter-like bacterial colonies were used for
804 DNA extraction. Lane 1, positive control containing DNA from CASJ002; Lanes 2-22,
805 samples collected from different fields in California; Lane 23, tomato DNA. The RhuM-
806 F/R primer pair targets the *rhuM* gene, the TomA-F1/R1 primer pair targets the *tomA*
807 gene, and the 16sR-F/R primer pair targets the 16SrRNA gene. Lane M, 1kb DNA
808 ladder, NEB™. **B)** Bacterial wilt symptoms induced by *C. michiganensis* strains
809 CASJ002, CA16_1068, CA16_1071, CA16_1073, CA16_1083, CA16_1085, and
810 CA16_1094 in tomato 14 days-post inoculation (dpi). Three-week-old tomato plants
811 (*Solanum lycopersicum* cv. Early Pak 7) were stabbed with a needle and inoculated with
812 a 5 µl bacterial suspension of 10^7 CFU/ml. **C)** Bacterial titers in infected tomato plant
813 materials reached similar levels to that of CASJ002 14 dpi. Three-week-old tomato
814 plants (*Solanum lycopersicum* cv. Early Pak 7) were stabbed with a needle and
815 inoculated with a 5 µl bacterial suspension of 10^7 CFU/ml. n = 6 plants per strain, error
816 bars = SD.

817

818

SUPPORTING INFORMATION LEGENDS:

Table S1. Bacterial strains used in this study. SN = strain number, Isolation = year of isolation, PCR = positive (+) indicates a band amplified with primer pairs targeting each locus. PDDCC = Culture Collection of the Plant Disease Division, CFBP = French Collection of Plant-associated Bacteria, and ATCC = American Type Culture Collection.

Table S2. *Clavibacter michiganensis* detection primers used in this study. Primers were tested on a panel of 110 bacterial strains (Table S1), including 75 *C. michiganensis* pathogens on tomato, 20 other *Clavibacter* strains, 6 *Pseudomonas,* and 9 *Xanthomonas* strains.

Table S3. Samples collected from counties in California with bacterial canker disease in 2016. Tomato samples exhibiting bacterial canker disease symptoms were collected from different counties and processed. Samples included fruits, leaves and stems. *C. michiganensis* was verified by mPCR, isolation on semi-selective media and pathogenicity assays.

Figure S1. *Clavibacter* strain collection and genome-wide comparison. A) Map showing the locations of *C. michiganensis* strains sequenced in this study. The circle size reflects the number of strains. **B)** Circos plots depicting genome-wide comparisons of the *C. michiganensis* strains NCPP382, NZ2541, and CASJ001. Chromosomes are depicted in green and orange (outer ring). Black bars depict unique genes, light green

bars depict orthologous genes, and the dark blue bar depicts % identity. Colored links connecting genomes are arbitrary.

Figure S2. Orthologous genes are distributed throughout the *Clavibacter* genome. Circos plots depicting the genome-wide distribution of orthologous genes (shown as blue lines) in the chromosome of *C. michiganensis* strain CASJ002. The large gap is the *chp/tomA* pathogenicity island.

Figure S3. **The limit of detection of the multiplex PCR with DNA extracted from bacterial cultures. A)** Purified CASJ002 DNA was diluted and one ul used for the multiplex PCR. The total bacterial DNA concentration in each multiplex PCR is indicated above each lane. **B)** CASJ002 bacterial cultures were subjected to crude DNA extraction with lysis buffer, DNA pellets were suspended in 20 ul of water, and one ul was used in the multiplex PCR. The estimated number of bacterial cells in each culture is indicated above each lane. Lane 1 = 1 kb DNA ladder.

Figure S4. Disease symptoms in tomato plant and fruit samples collected from a 2016 outbreak of bacterial canker in central California.

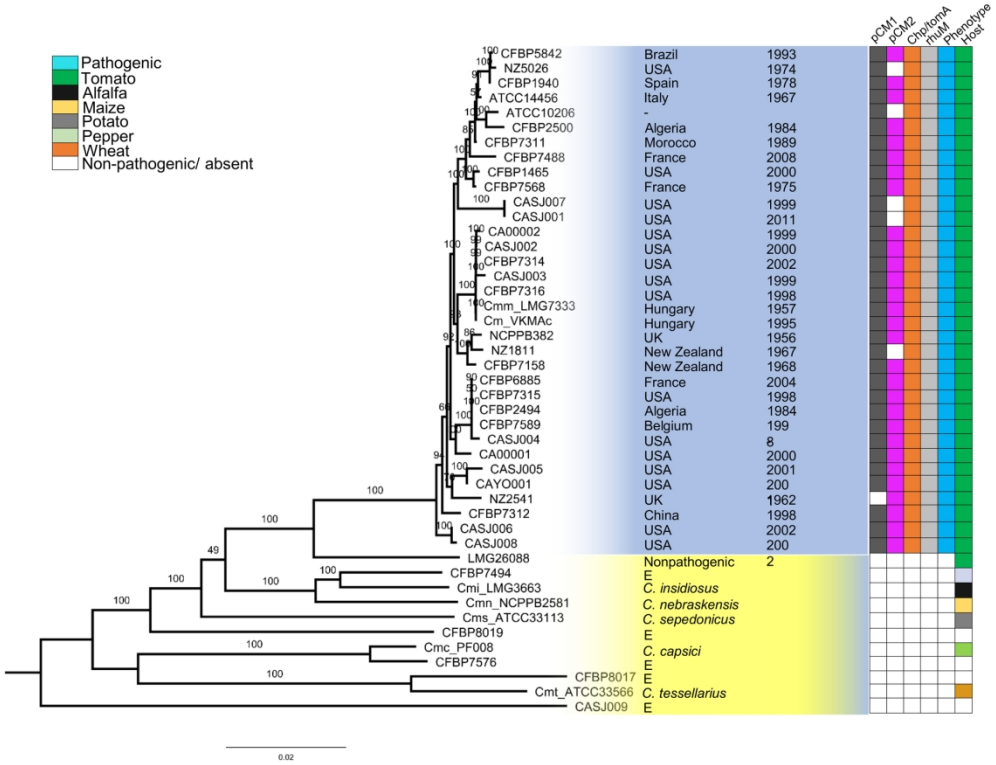

Figure 1. Genome assisted phylogeny of the genus Clavibacter. A total of 1,258 orthologous genes from each strain were concatenated. A maximum likelihood approach was used to generate the phylogeny with 1000 bootstrap replicates. Bootstrap values are indicated at each node. Columns indicate the presence of the pCM1 plasmid, the pCM2 plasmid, the chp/tomA genomic island, and RhuM. Phenotype = blue indicates pathogenicity on tomato, Host = plants bacteria were isolated from if known, E = endophyte of tomato.

175x134mm (300 x 300 DPI)

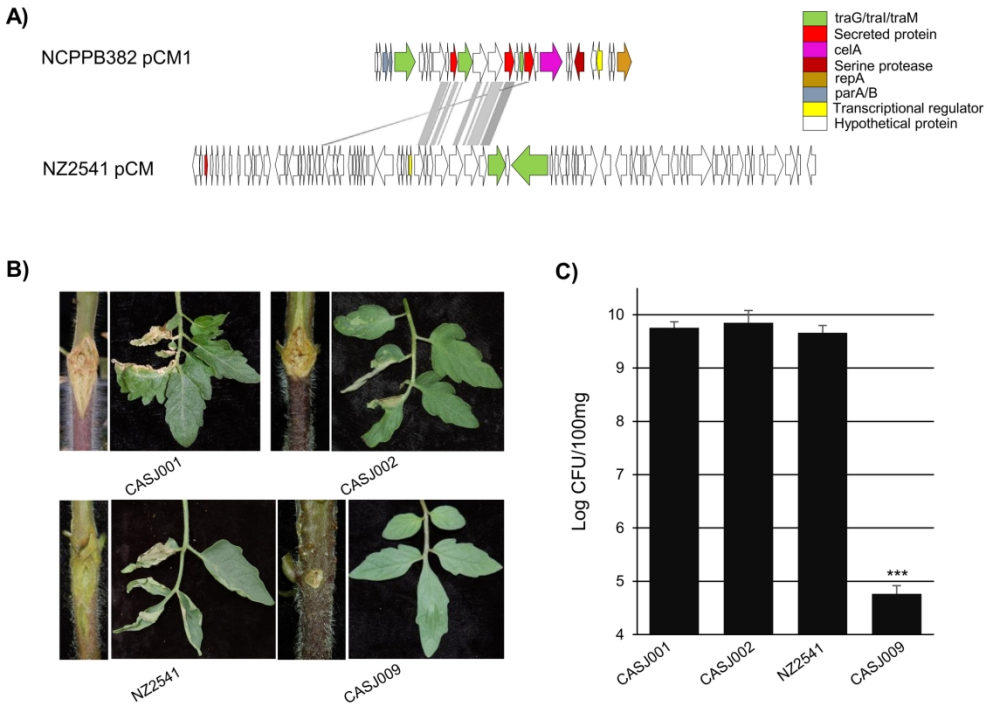

Figure 2. C. michiganensis strain NZ2541 lacks the pCM1 plasmid and celA but is pathogenic on tomato. A) Comparison of NCPPB382's pCM1 plasmid with the new plasmid found in NZ2541, pCM3. White = hypothetical proteins, red = secreted proteins, yellow = transcriptional regulator, violet = conjugal transfer protein, pink = celA. B) Representative bacterial wilt and canker symptoms induced by C. michiganensis strain in tomato 7 days post-inoculation (dpi). C) Bacterial titers in infected tomato stems 7 dpi. Three-week-old tomato plants (Solanum lycopersicum cv. Early Pak 7) were inoculated by stabbing tomato stems with a needle and delivering 5 µl of at 107 CFU/ml of bacteria. Bacterial titers were determined by grinding sterilized stem segments 7 dpi and plating on D2 semi-selective media. Statistical differences were detected by a two-tailed t-test compared and; asterisks indicate P = 0.001.

186x133mm (300 x 300 DPI)

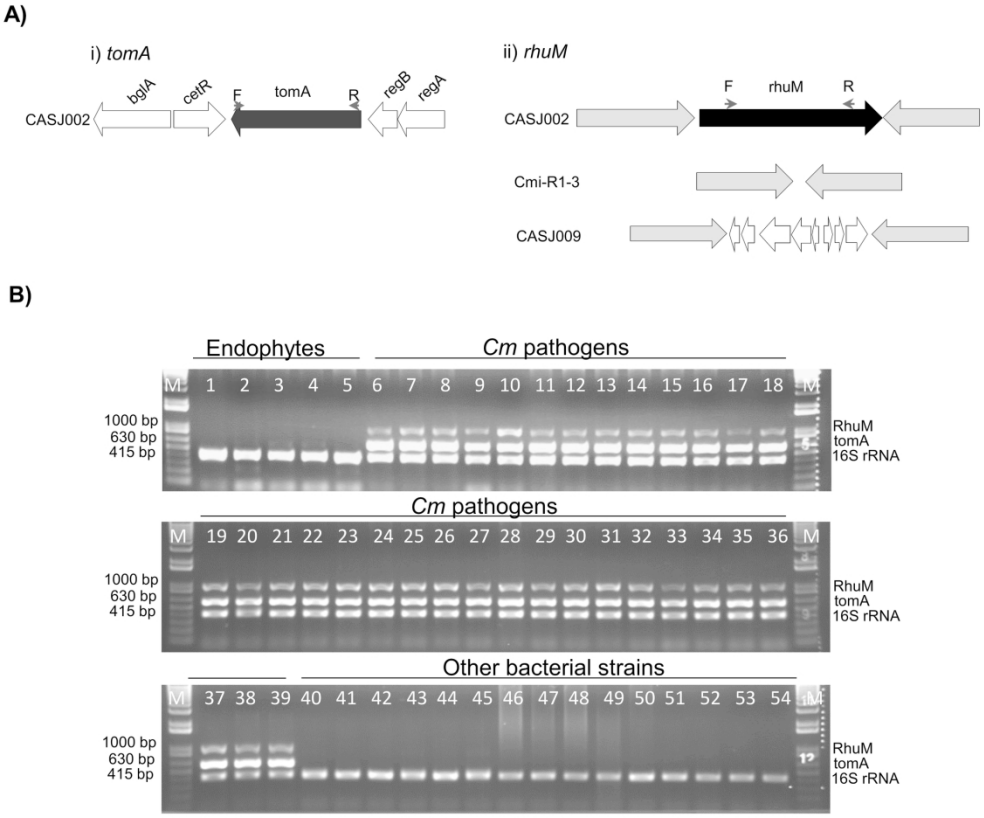

Figure 3. C. michiganensis primer design and multiplex PCR. A) The genetic architecture of flanking genes surrounding the tomA and rhuM loci. The chp/tomA cluster, including tomA, is only present in pathogenic C. michiganensis. RhuM is only present in C. michiganensis tomato pathogens. Primers used to amplify each gene are shown in grey. B) Agarose gel electrophoresis of multiplex PCR assay products. Total genomic DNA was used as a template with the RhuM-F/R primers targeting the rhuM gene, the TomA-F1/R1 primers targeting the tomA gene, and the 16sR-F/R primers targeting 16S rRNA in either bacteria or plants. Lanes 1-5, endophytic Clavibacter strains (CASJ009, CFBP8017, CFBP8019, CFBP7576, and CFBP7494); lanes 6-39, pathogenic C. michiganensis (CFBP7311, CFBP1940, CFBP5842, CFBP7568, CFBP7488, CFBP2500, CFBP7158, CFBP7312, CFBP7316, CFBP7314, CFBP6885, CFBP7315, CFBP7589, NCPPB382, ATCC14456, ATCC10206, NZ2541, NZ1811, NZ5026, CASJ001, CASJ002, CASJ003, CASJ004, CASJ005, CASJ008, CA00001, CAYO001, AZ-28, NZ2550, 285D, 12B, 13C1, Cm4, and Cm5); lanes 40-43, C. sepedonicus (NZ2535, SB109, INM-1, and BRR7); lane 44, C. insidiosus (CIC266); lanes 45-46, C. tessellarius (CIC021,CIC022); lanes 47-48, Xanthomonas euvesicatoria (RL677, HB1); lanes 49-51 Pseudomonas syringae pv. tomato (PstA9, Pst18, and Pst838-8); lanes 52-54, E. coli, Candidatus Liberibacter asiaticus HHCA, and uninfected Tomato DNA. Lane M = 1kb DNA ladder.

184x155mm (300 x 300 DPI)

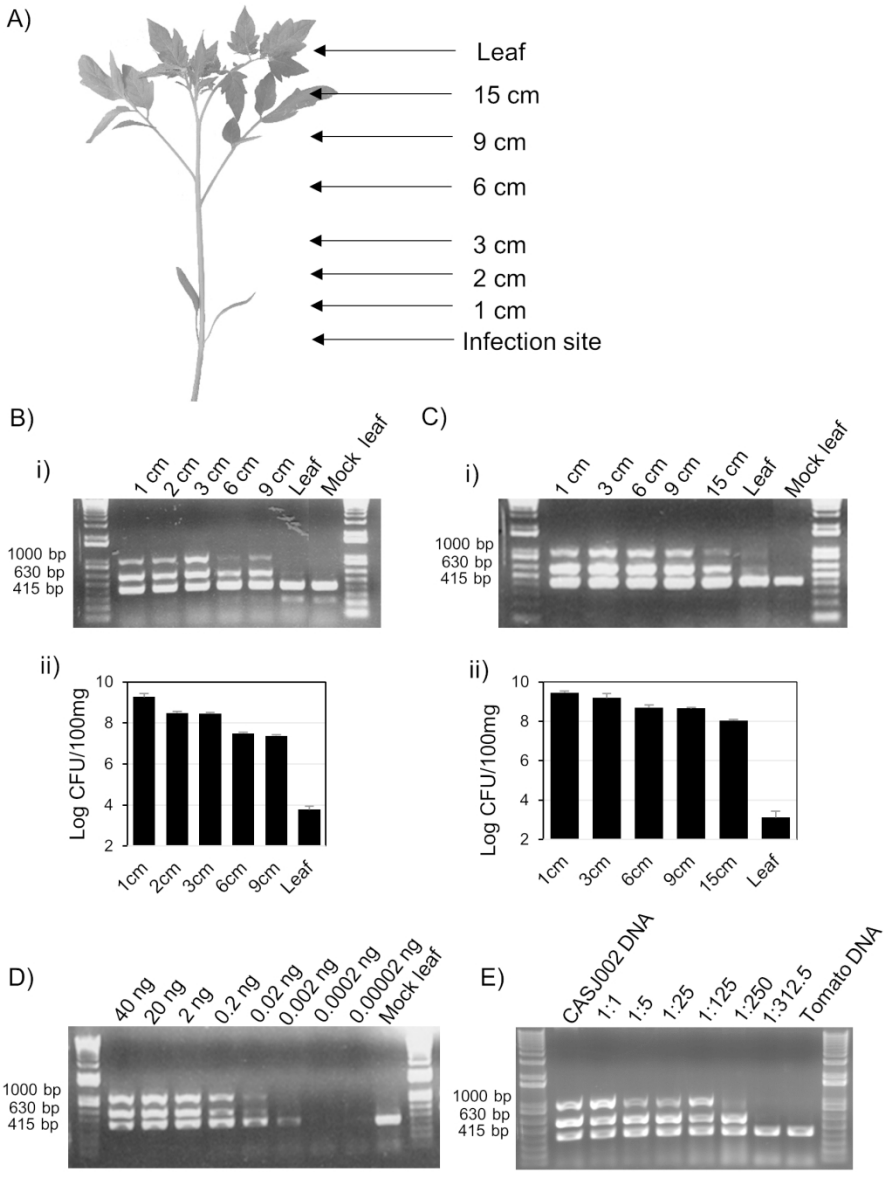

Figure 4. Detection of C. michiganensis in infected plant materials. A). Three-week-old tomato plants (Solanum lycopersicum cv. Early Pak 7) were stabbed with a needle and inoculated with 5 µl of 107 CFU/ml suspension of CASJ002. The diagram indicates points of sample collection above the C. michiganensis inoculation site. B-i, ii) Detection of C. michiganensis by multiplex PCR and bacterial titers in inoculated tomato plants 7 days post-inoculation (dpi). DNA was extracted from 100 mg of plant tissue with the CTAB method, re-suspended in 50 µl of water, and 2 µl of DNA was used per multiplex PCR reaction. C-i, ii) Detection of C. michiganensis by multiplex PCR and bacterial titers in CASJ002 inoculated tomato plants at 14 dpi. Bacterial titers, DNA extraction and multiplex PCR were performed as described in B. D) The detection limit for CASJ002 after CTAB DNA extraction from tomato stems. DNA was extracted from plant samples 9 cm above the site of inoculation at 7 dpi. DNA samples were diluted and the total DNA concentration in each multiplex PCR reaction is indicated above each lane. E) Detection of C. michiganensis by multiplex PCR in serial dilutions of CASJ002 DNA (50ng) and healthy tomato plant DNA (50 ng) mixed in different ratios using 100 ng/ul in a 25 µl PCR reaction. Numbers indicate bacterial:plant DNA ratios. Lane M, size marker (1kb DNA ladder).

158x204mm (300 x 300 DPI)

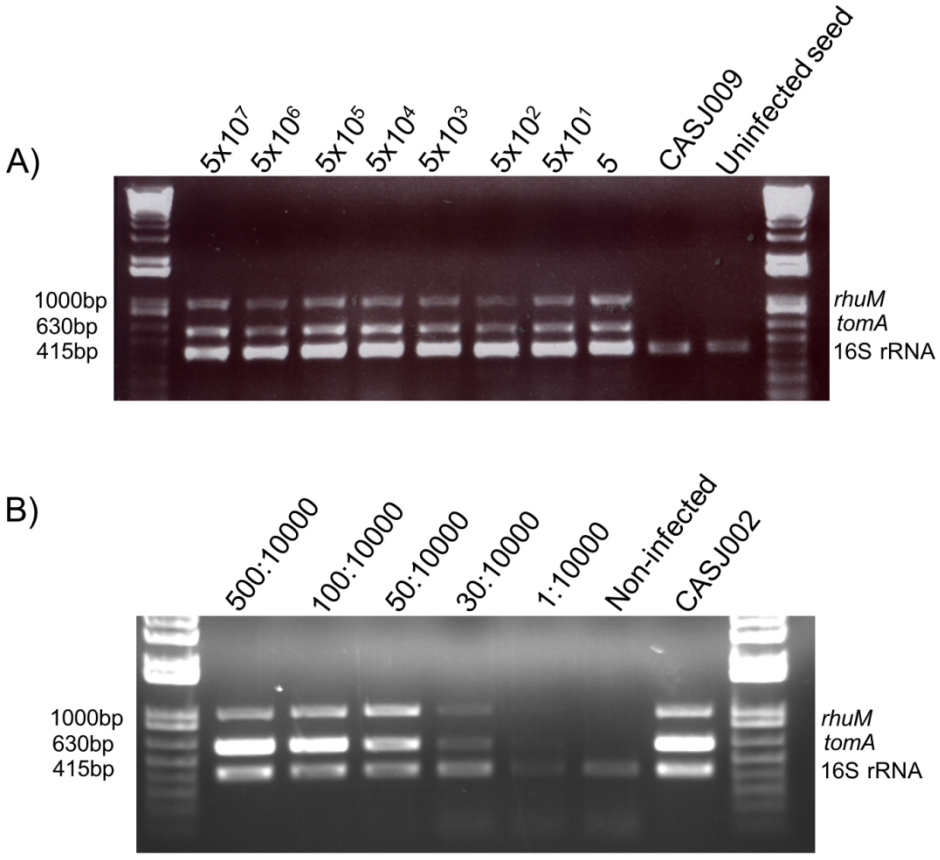

Figure 5. Detection of C. michiganensis in seeds. A) C. michiganensis multiplex PCR detection from contaminated tomato seed. Tomato seeds were contaminated with CASJ002 at the indicated concentrations. DNA was directly extracted from 10 seeds, the pellet re-supended in 50 ul, and 2 ul of purified DNA was used in each multiplex PCR reaction. B) Detection of C. michiganensis by multiplex PCR after spiking contaminated seed with clean seed lots consisting of 10,000 seed. Seeds infested with 5x107 CASJ002 were mixed with clean seeds in different ratios followed by DNA extraction and multiplex PCR. Seeds were incubated with 120 ml of PBS, crushed in a stomacher, and 5-10 ml was used for DNA extraction. The pellet was re-suspended in 100-200 ul and 2-4 ul of purified DNA was used in each PCR reaction. Infested: clean seed ratios are indicated above each lane. Lane1, 1kb DNA ladder.

123x111mm (300 x 300 DPI)

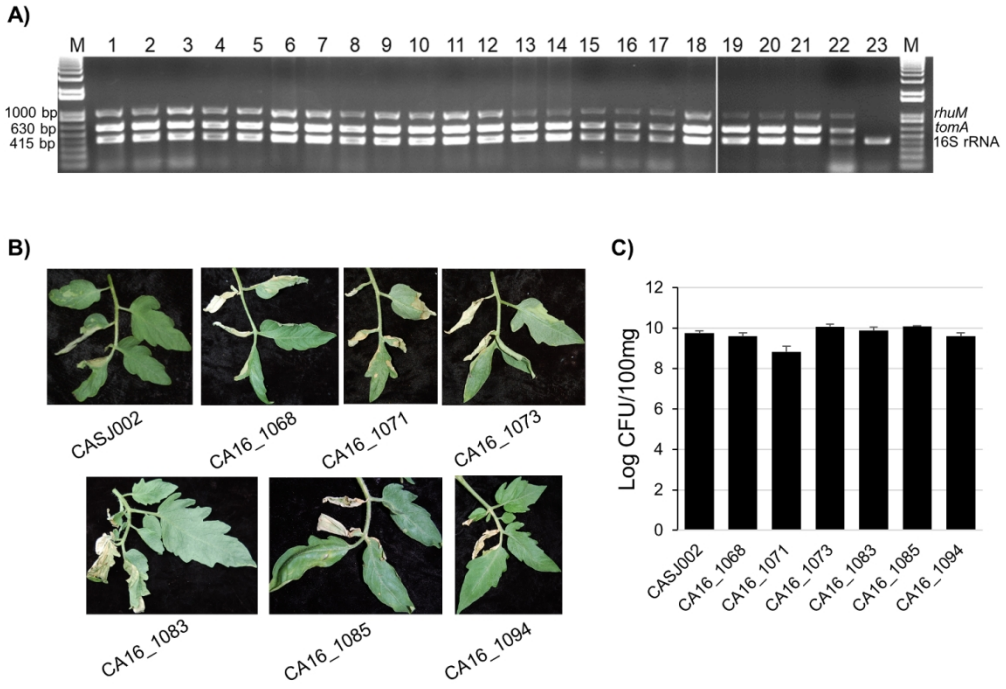

Figure 6. Multiplex PCR confirmed the presence of C. michiganensis in samples collected from commercial production fields. A) Detection of C. michiganensis by multiplex from field samples. Purified Clavibacter-like bacterial colonies were used for DNA extraction. Lane 1, positive control containing DNA from CASJ002; Lanes 2-22, samples collected from different fields in California; Lane 23, tomato DNA. The RhuM-F/R primer pair targets the rhuM gene, the TomA-F1/R1 primer pair targets the tomA gene, and the 16sR-F/R primer pair targets the 16SrRNA gene. Lane M, 1kb DNA ladder, NEB™. B) Bacterial wilt symptoms induced by C. michiganensis strains CASJ002, CA16_1068, CA16_1071, CA16_1073, CA16_1083, CA16_1085, and CA16_1094 in tomato 14 days-post inoculation (dpi). Three-week-old tomato plants (Solanum lycopersicum cv. Early Pak 7) were stabbed with a needle and inoculated with a 5 µl bacterial suspension of 107 CFU/ml. C) Bacterial titers in infected tomato plant materials reached similar levels to that of CASJ002 14 dpi. Three-week-old tomato plants (Solanum lycopersicum cv. Early Pak 7) were stabbed with a needle and inoculated with a 5 µl bacterial suspension of 107 CFU/ml. n = 6 plants per strain, error bars = SD.

127x89mm (300 x 300 DPI)

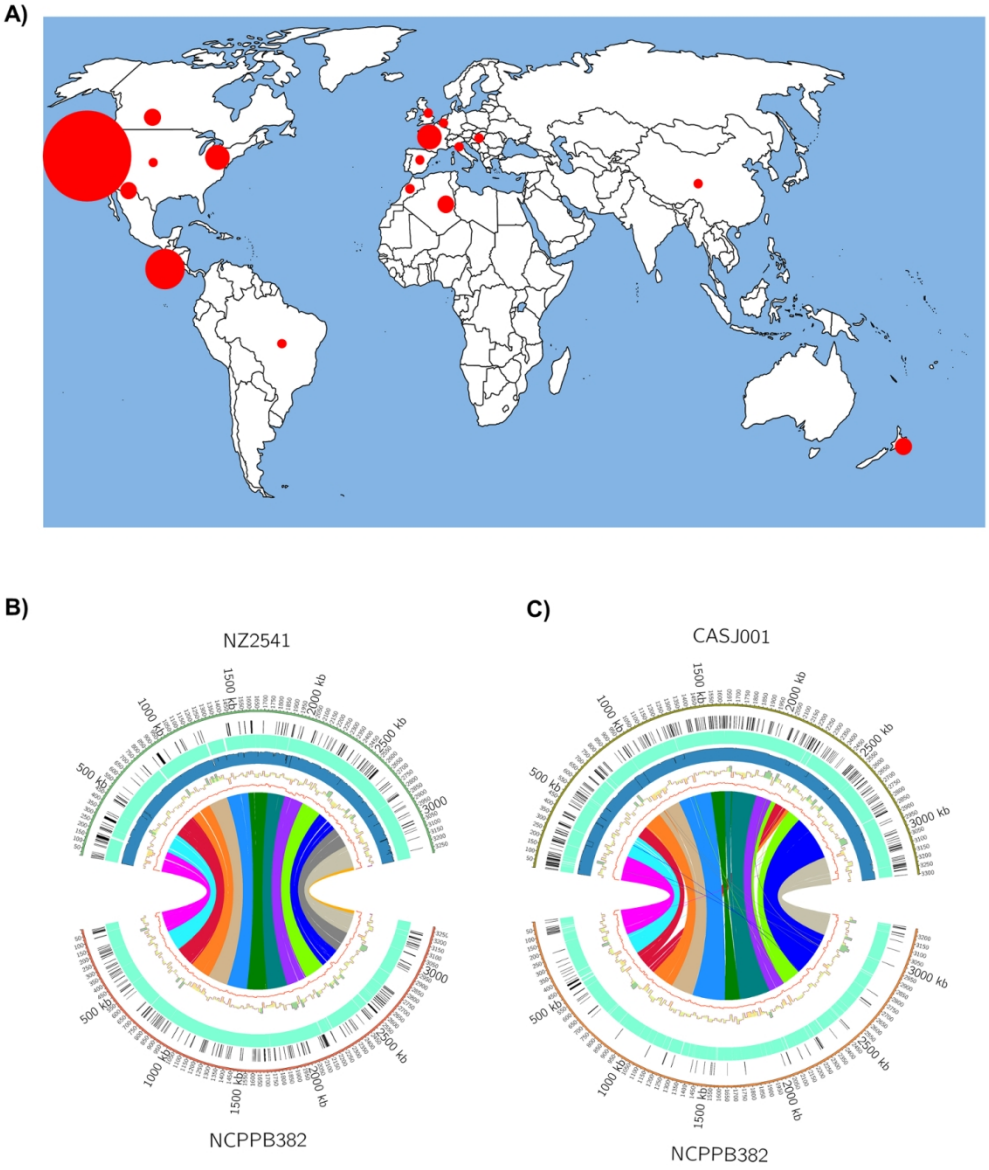

Figure S1. Clavibacter strain collection and genome-wide comparison. A) Map showing the locations of C. michiganensis strains sequenced in this study. The circle size reflects the number of strains. B) Circos plots depicting genome-wide comparisons of the C. michiganensis strains NCPP382, NZ2541, and CASJ001. Chromosomes are depicted in green and orange (outer ring). Black bars depict unique genes, light green bars depict orthologous genes, and the dark blue bar depicts % identity. Colored links connecting genomes are arbitrary.

127x150mm (300 x 300 DPI)

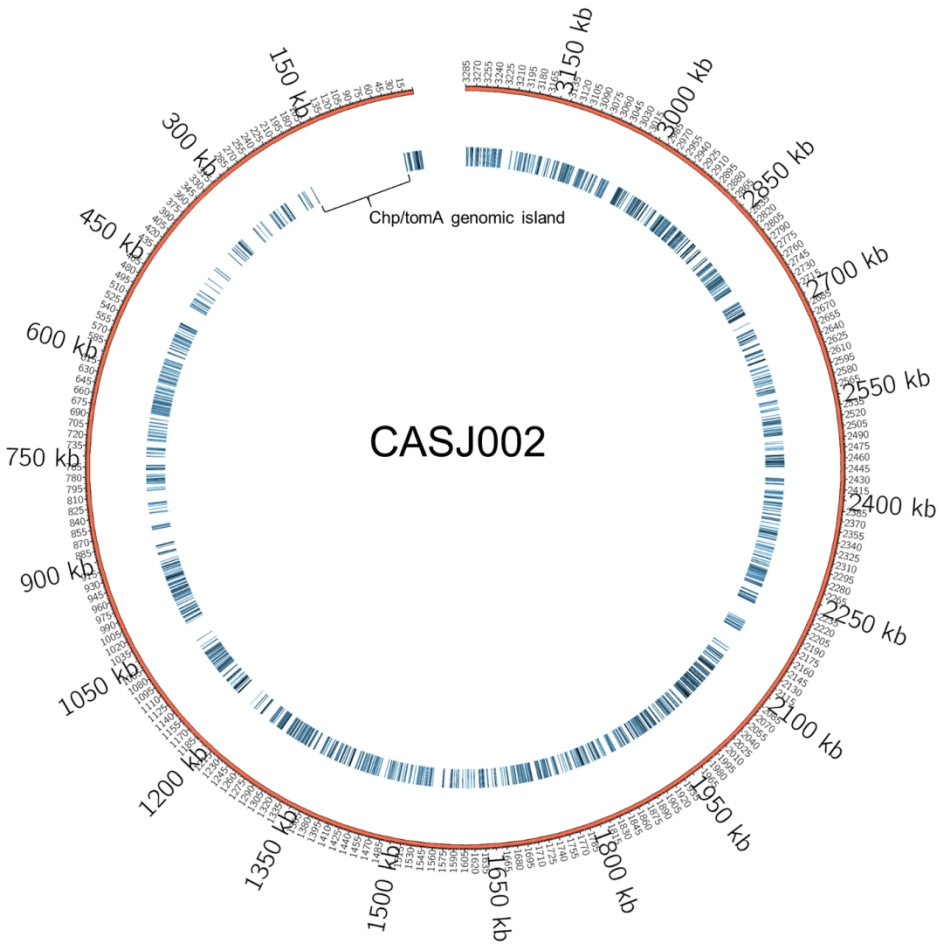

Figure S2. Orthologous genes are distributed throughout the Clavibacter genome. Circos plots depicting the genome-wide distribution of orthologous genes (shown as blue lines) in the chromosome of C. michiganensis strain CASJ002. The large gap is the chp/tomA pathogenicity island.

127x122mm (300 x 300 DPI)

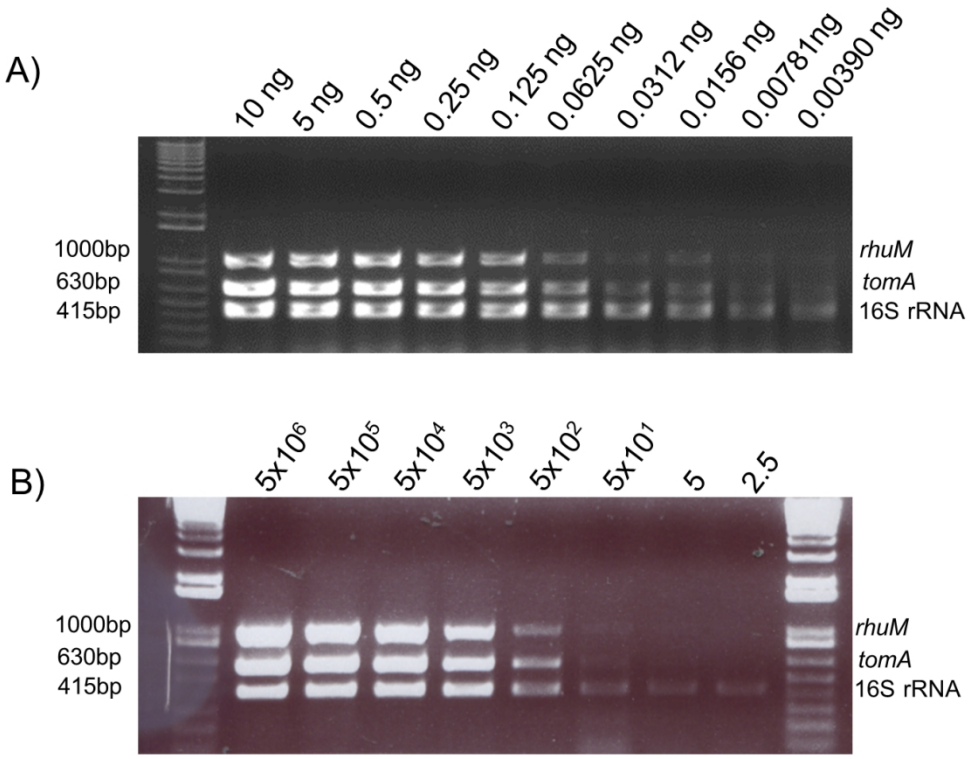

Figure S3. The limit of detection of the multiplex PCR with DNA extracted from bacterial cultures. A) Purified CASJ002 DNA was diluted and one ul used for the multiplex PCR. The total bacterial DNA concentration in each multiplex PCR is indicated above each lane. B) CASJ002 bacterial cultures were subjected to crude DNA extraction with lysis buffer, DNA pellets were suspended in 20 ul of water, and one ul was used in the multiplex PCR. The estimated number of bacterial cells in each culture is indicated above each lane. Lane 1 = 1 kb DNA ladder.

118x96mm (300 x 300 DPI)

Figure S4. Disease symptoms in tomato plant and fruit samples collected from a 2016 outbreak of bacterial canker in central California.

135x41mm (300 x 300 DPI)

TABLE S2. *Clavibacter michiganensis* detection primers used in this study. Primers were tested on a panel of 110 bacterial strains (Table S1), including 75 *C. michiganensis* pathogens on tomato, 20 other *Clavibacter* strains, 6 *Pseudomonas,* and 9 *Xanthomonas* strains.

Sequence (5'-3')	Target	Amplicon size (bp)	*Cmm* (75)	Other bacterial strains (35)	Reference
PFC3: GGTACGAAGTTCGAGACGAC **PFC5:** TGTAGCGGTGAGTCGTGGTGA	*celA*	533	74	0	Kleitman et al. 2008
P5: GCGAATAAGCCCATATCAA **P6:** CGTCAGGAGGTCGCTAATA	*Pat-1*	1474	70	0	Kleitman et al. 2008
ppaAF: CATGATATTGGTGGGGAAAG **ppaAR:** CCCCGTCTTTGCAAGACC	*ppaA*	768	73	2	Kleitman et al. 2008
tomAF: CGAACTCGACCAGGTTCTCG **tomAR:** GGTCTCACGATCGGATCC	*tomA*	320	75	0	Kleitman et al. 2008
celAF: ATGGCTTCCCTACGATCC **celAR:** ACAGGGTAGAAGCGGGAGG	*celA*	2193	74	0	Jahr et al. et al 2000
CMM3: CCTCGTGAGTGCCGGGAACGTATCC **CMM4:** CCACGGTGGTTGATGCTCGCGAGAT	*ppaJ*	645	72	0	Santos et al.1997
CMM5: GCGAATAAGCCCATATCAA **CMM6:** CGTCAGGAGGTCGCTAATA	*pat-1*	614	70	0	Drier et al. 1995
PSA-F: TCATTGGTCAATTCTGTCTCCC **PSA-R:** TACTGAGATGTTTCACTTCCCC	16S-23S	290	75	5	Pastrik and Rainey, 1999
TomA-F1: ATGAAGAGCTTCGCGTCCG **TomA-F2:** GAGAACACTGACATCCGCAG	*tomA*	630	75	0	This study
RhuM-F: GGGTCGGTTCATCCTGTA **RhuM-R:** CTTCGGGAGGTTCTCCTGT	*rhuM*	1000	75	0	This study
16sR-F: TTGCGGGACTTAACCCAAC **16sR-R:** AGCGGTGAAATGCGCAGA	16s RNA	415	75	0	This study

Dreier, J., Bermpohl, A., & Eichenlaub, R. (1995). Southern hybridization and PCR for specific detection of phytopathogenic *Clavibacter michiganensis* subsp. *michiganensis*. Phytopathology, 85(4), 462–468.

Santos, M. S., Cruz, L., Norskov, P., & Rasmussen, O. F. (1997). A rapid and sensitive detection of in tomato seeds by polymerase chain reaction. Seed Science and Technology, 25, 581–584.

Jahr, H., Dreier, J., Meletzus, D., Bahro, R., & Eichenlaub, R. (2000). The endo-β-1,4-glucanase CelA of *Clavibacter michiganensis* subsp. *michiganensis* is a pathogenicity determinant required for induction of bacterial wilt of tomato. Molecular Plant-Microbe Interactions, 13(7), 703–714.

Kleitman, F., Barash, I., Burger, A., Iraki, N., Falah, Y., Sessa, G., Weinthal, D., Chalupowica, L., Gartemann, K. H., Eichenlaub, R., & Manulis-Sasson, S. (2008). Characterization of a *Clavibacter michiganensis* subsp. *michiganensis* population in Israel. European Journal of Plant Pathology, 121, 463–475.

Pastrik, K.H. and Rainey, F.A. 1999. Identification and differentiation of *Clavibacter michiganensis* subspecies by polymerase chain reaction-based techniques. J. Phytopathol. 147, 687–693.