A host shift as the origin of tomato bacterial canker caused by *Clavibacter michiganensis*

Alan G. Yañez-Olvera[1,3], Ámbar G. Gómez-Díaz[1], Nelly Selem-Mojica[1,4], Lorena Rodríguez-Orduña[1,5], José Pablo Lara-Ávila[1,6], Vanina Varni[1,7], Florencia Alcoba[2], Valentina Croce[2], María Inés Siri[2], Clavibacter Consortium, Francisco Barona-Gómez[1,3,*]

[1] Evolution of Metabolic Diversity Laboratory, Unidad de Genómica Avanzada (Langebio), Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional (Cinvestav-IPN), Irapuato, Guanajuato, Mexico

[2] Laboratorio de Microbiología Molecular, Departamento de Biociencias, Facultad de Química, Universidad de la República, Montevideo, Uruguay

[3] Institute of Biology, Leiden University, Leiden, the Netherlands

[4] Current address: Centro de Ciencias Matemáticas, UNAM, Morelia, Michoacán, Mexico.

[5] Current address: Tecnológico de Monterrey, Escuela de Ingeniería y Ciencias, Monterrey, Mexico.

[6] Current address: Facultad de Agronomía y Veterinaria, Universidad Autónoma de San Luis Potosí, San Luis Potosí, Mexico.

[7] Current address: Laboratorio de Genómica Computacional, Universidad Nacional de Luján, Buenos Aires, Argentina.

*Corresponding author: f.barona.gomez@biology.leidenuniv.nl

Keywords. *Clavibacter michiganensis,* phylogenomics, pangenomics, host shift, bacterial canker, tomato.

Abstract

Clavibacter, a plant endophytic actinobacterial genus, includes phytopathogens with devasting effects on several crops. *C. michiganensis*, the seed-borne and causal agent of bacterial canker in tomato, is arguably the most notorious species of the genus. Yet, its origin and natural reservoirs remain elusive. Moreover, *C. michiganensis* populations show different genetic pathogenicity profiles with equally unpredictable plant disease outcomes. To tackle these uncertainties, here we analyze genomic data generated during a decade-long survey of *Clavibacter* in wild and commercial tomato cultivars, providing evolutionary insights that informed on the pathogenicity of this phytopathogen. Unexpectedly, our phylogeny situate the last common ancestor of *C. michiganensis* next to *Clavibacter* isolates from grasses rather than to the sole strain we could isolate from wild tomato, which is closer to *C. capsici* associated with pepper. Pathogenicity profiling of selected *C. michiganensis* isolates, together with *C. phaseoli* and *C. californiensis* as sister taxa of the grass clade, and the newly isolated *C. capsici* from wild tomato, was found to be congruent with the proposed phylogenetic relationships. Furthermore, we identified gene enrichment after an evolutionary bottleneck leading to the appearance of *C. michiganesis*, including known pathogenicity factors but also hitherto unnoticed genes with such potential, *i.e.,* nutrient acquisition and specialized metabolite metabolic gene clusters. The holistic perspective provided by our long-term and in-depth analyses hints towards a host shift event as the origin of the causative agent of bacterial canker in tomato, leading to a complex of *C. michiganensis* with pathogenicity factors that remain to be characterized.

Introduction

Many aspects of human activity are related with the emergence of new pathogens (Sabin et al., 2020). Phenomena like climate change, human invasion of ecosystems and urbanization, can break ecological barriers and create paths that potential pathogens can use to reach new niches (Zhang et al., 2022). Agriculture is no exception: monocultures, as well as high-density cultivation, have been linked to the emergence of new phytopathogens (Stukenbrock & McDonald, 2008), as these practices impose new selective pressures that may favor virulent lineages (McDonald & Stukenbrock, 2016). The introduction of crops to novel environments is another factor that may cause the emergence of pathogens (Weisberg et al., 2021), as this can be transmitted from wild to domesticated hosts, even from unrelated species (McCann, 2020). During this process host shifts take place (Thines, 2019), representing the origin of an emerging disease.

Bacteria of the *Clavibacter* genus, which belong to the phylum Actinomycetota, are mainly known for their phytopathogenic members capable of generating severe diseases in crops of great economic importance (Eichenlaub et al., 2006; Mansfield et al., 2012). The *Clavibacter* phylogeny originally included six subclades referred to as subspecies, which have now been recognized as species (Li et al., 2018) with several new species being proposed (Arizala et al., 2022; Osdaghi et al., 2020). Amongst these, *C. michiganensis* has received a great deal of attention as it is responsible for bacterial canker disease in tomato, *Solanum lycorpersicum* (Sen et al., 2015), with devastating consequences (Poysa, 1993; Mansfield et al., 2012). Even though it was first isolated more than a century ago from tomato crops in Michigan, USA (Smith, 1910), molecular markers are the sole means to tackle the infection, yet reproducibility has been shown to be an issue (Thapa et al., 2020; Yasuhara-Bell et al., 2014). The latter may reflect the fact that genetically diverse *C. michiganensis* populations exist around the globe (Kleitman et al., 2008; Tancos et al., 2015; Valenzuela et al., 2021; Wassermann et al.,

2020), and the fact that the origins and natural reservoirs of this phytopathogen remain unknown.

It was not until recently that researchers began to describe genetic features that make this bacterium pathogenic. For instance, Meletzus and co-workers (Meletzus et al., 1993) found evidence that plasmids harbored by *C. michiganensis* contained sequences that encode virulence factors: *celA* on the pCM1 plasmid, which encodes a endo-β-1,4-glucanase (Jahr et al., 2000), and *pat-1* on pCM2, which encodes a serine protease (Dreier et al., 1997). Genome sequencing of the reference NCPPB 382 strain (Stork et al., 2008) allowed the identification of a pathogenicity island (PAI) present in the chromosome. Further genomic studies have shown that *C. michiganensis* has unique features that could explain its pathogenic character in tomato. For instance, Thapa and co-workers (Thapa et al., 2017) showed that certain CAZymes (Carbohydrate-Active Enzymes) families involved in cellulose and hemicellulose degradation are abundant in the *C. michiganensis* genome. Although these previously identified pathogenicity factors have been unambiguously shown to be related to the development of the disease, there is also evidence that some of these factors are absent from strains capable of causing symptoms of the disease (Oh et al., 2022; Tancos et al., 2015) and that virulent strains can infect tomato plants without leading to symptoms (Gitaitis, 1991). Together, these observations suggest that other factors are required for triggering the disease and symptom development (Sharabani et al., 2013, 2014).

Comparative genomics of *C. michiganensis* and other members of its genus has shown that while each species has acquired unique characteristics (Tambong, 2017), there are several features linked to pathogenicity shared between *C. michiganensis* and other *Clavibacter* strains, including cases without background or an antecedent of pathogenicity (Osdaghi et al., 2020). Nevertheless, genomic comparisons have overlooked the relationship between their hosts from which they were isolated and the genomic features that these bacteria may have acquired in their path to adapt to their hosts and develop a pathogenicity lifestyle. To tackle these limitations, here we perform a comprehensive phylogenomics and pangenomic analysis of the *Clavibacter* genus, with an emphasis on *C.*

michiganensis strains isolated throughout a decade in geographically distantly and unrelated greenhouses in North (Mexico) and South (Uruguay) America, as well as Europe (the Netherlands). We provide evidence for *C. michiganensis* species sharing a common ancestor with isolates from different grasses, and that the species has undergone an evolutionary bottleneck. Our analysis revealed genetic markers, with functional implications *in planta*, that evidence a host shift, from grasses to tomato. We provide an example on how to track the origin of emerging plant pathogens and pave the way to further functionally characterize the pathogenicity of *C. michiganensis* at the population level in line with its dual symptomatic and asymptomatic behavior.

Methods

Bacterial strain isolation and identification. Tomato (*Solanum lycopersicum*) plants with symptoms but also asymptomatic were collected during 2010 to 2020 from several geographically distant sites in Mexico, mainly from high-tech greenhouses (Aguascalientes, Colima, Guanajuato, Michoacán, Nuevo León, Querétaro, San Luis Potosí and Zacatecas), and from two wild populations (Jalisco and Guanajuato). Samples consisted of stem, leaves, and fruit when available. Tissue was cut into slices of approximately 0.5 cm^2 of surface area and placed directly into petri dishes with CMM1 semi-selective media (for 1 liter: 10.0 g saccharose 1.2 g Tris base, 250 mg $MgSO_4 \cdot 7H_2O$, 5.0 g LiCl, 2.0 g yeast extract, 1.0 g NH_4Cl, 4.0 g casamino acids and 15.0 g agar; supplemented with 28 mg/l nalidixic acid, 10 mg/l polymyxin B sulfate and 200 mg/l cycloheximide). CMM1 Petri dishes were incubated at 28°C and growth monitored for 24 to 72 hrs. Grown colonies were selected based on *C. michiganensis* reported morphology (EPPO, 2016) and isolated and cultured on individual petri dishes with CMM1 media. LB liquid cultures were prepared for each isolate for DNA extraction, carried out with alkaline lysis, phenol-chloroform extraction, and ethanol precipitation. *C. michiganensis* isolates were identified by PCR using the *clvF* marker gene (Yasuhara-Bell et al., 2014). *clvF* PCR negative isolates were identified by Sanger sequencing of the PCR amplified 16S rRNA gene. A total of 550 plant specimens, leading to a strain collection of around 151 *Clavibacter* positive isolates, were obtained (148 *C. michiganensis* confirmed strains).

Genome sequencing and database.

DNA from the 151 *Clavibacter* positive strains was extracted, as previously stated, for whole genome sequencing using Illumina paired-end MiSeq or NextSeq platforms. Read quality was assessed with FastQC. Poor quality sequences were trimmed with Trimmomatic (Bolger et al., 2014), with values adjusted for each read set. *De novo* assembly was performed with the smart and auto assembly strategies in PATRIC (Davis et al., 2020). Seven representative genomes of the *C. michiganensis* strains isolated in Mexico from commercial cultivars were chosen

based on independent phylogenomic and pangenome analysis made at species level followed by a selection of those strains that best represented the genomic diversity according to the gene families present in their respective clades (data unpublished).These selected genomes and 58 publicly available genomes were used to construct an *ad hoc* genomic database (DB) which was complemented with *C. michiganensis* genomes obtained from strains isolated from tomato cultivars with symptoms from other parts of the world: 3 from Uruguay and 1 from the Netherlands. From the total of 69 *Clavibacter* genomes that comprised the complete *Clavibacter* genus DB, 32 corresponded to *C. michiganensis* sequences. All genome assemblies comply with quality metric thresholds: ≥25000 for the N50 and ≤36 for L50. The RAST tool kit was used for gene calling and function prediction (Brettin et al., 2015), complemented with the Pfam database (Mistry et al., 2021).

Clavibacter *phylogeny and pangenome analysis*

Single copy core genes in the previous DB were identified using Anvi'o (v7.0) pangenome analysis tool (Eren et al., 2021). Identical genes were filtered out based on the functional homogeneity index (<0.999). A total of 1231 genes were selected. Aligned and concatenated amino acid sequences were obtained with *anvi-get-sequences-for-gene-clusters*. The concatenated sequences were used to infer phylogenetic relationships by generating phylogenetic trees with IQ-TREE 2 (Minh et al., 2020). Branch support was assessed using the ultrafast bootstrap approximation from UFBoot (Hoang et al., 2018) with 1000 replicates. Best-fit substitution model was determined using ModelFinder (Kalyaanamoorthy et al., 2017) restricting the testing procedure to the WAG (Whelan & Goldman, 2001) and LG (Le & Gascuel, 2008) models. Branches of the *Clavibacter* genus phylogeny were ordered using a family-level reference tree, which included 15 *Clavibacter* strains and *Rathayibacter toxicus FH232* as a root (**Supp. Fig. 1**). A pairwise Average Nucleotide Identity analysis of all the genomes was performed with *anvi-compute-genome-similarity* using PyANI (Pritchard et al., 2015) with BLAST and default values. The DB was analyzed using the program *anvi-pangenome* from

Anvi'o, using default parameters unless stated otherwise. Briefly, the program calculated the similarity between amino acid gene sequences with BLASTp (Altschul et al., 1997), then resolved gene families with the MCL algorithm (Van Dongen & Abreu-Goodger, 2012) using an inflation value of 10. Eren et al. (2021) refer to *gene clusters*, which we call *gene families* in this manuscript, to avoid confusion when referring to Biosynthetic Gene Clusters (BGCs).

Identification of genomic evolutionary signals.

Gene family enrichment analysis was performed at the genus-level with the cognate genome DB using anvi'o's *anvi-compute-functional-enrichment* program and the *–include-gc-identity-as-function* option. Briefly, this program groups the genomes in the dataset according to a categorical variable, in our case membership to the so-called "*C. michiganensis* broad clade", and then determines which gene families are enriched in the genomes within our group of interest and absent, or nearly absent, in the rest of the genomes. The statistical approach to determine the enrichment was described elsewhere (Shaiber et al., 2020). Parallelly, we looked for phylogenetic signals in the occurrence of the identified gene families with the D measure (Fritz & Purvis, 2010) using caper (v. 1.0.1) R package (Orme et al., 2013). D was estimated for each gene family using the *Clavibacter* genus phylogeny and the presence and absence of the gene families in each genome as the binary trait to evaluate. Gene families were considered enriched when their enrichment score was $\geqslant$50, their q- adjusted value from the enrichment analysis was < 1e-10 and the value of D was < 0. For the identification of evolutionary and functionally informative loci, genes from the reference strain NCPPB 382 occurring in the enriched gene families were extracted using Anvi'o pangenome analysis tool. Since this is a closed genome, their annotation order corresponds to their position in the chromosome, hence genes were ordered accordingly and checked for contiguity. Only genes with at least three enriched neighboring genes were considered for further inspection. Genes were still considered neighbors even when there was a gap of maximum two non-enriched genes between them. Genes within the pathogenicity island (PAI) were identified

according to their location relative to the genes that limit this region as reported by Gartemann et al. (2008).

Evolutionary genome mining of natural products.

Given the results obtained during the identification of evolutionary genomic traits, we focused on the prediction and analysis of BGCs encoding for Ribosomally synthesized and Post-translationally modified Peptides (RiPPs). RiPP type BGCs were searched in the genome DB using antiSMASH v6.0 (Blin et al., 2021), including RRE-Finder (Kloosterman et al., 2020) to improve RiPP BGC identification. Identified RiPP BGCs were grouped into similarity networks using BiG-SCAPE (Navarro-Muñoz et al., 2020) employing a cut off value of 5.0 and compared against version 3.0 of MiBIG database (Terlouw et al., 2022). Selected BGCs were analyzed further with CORASON to solve their evolutionary relationships at the whole BGC-level (Navarro-Muñoz et al., 2020). Predicted structures were based on previously reported literature (Holtsmark et al., 2006; Wiebach et al., 2018) and are provided only as a proxy to chemical classes involved, but remain to be experimentally confirmed. CORASON (Navarro-Muñoz et al., 2020) was used to explore the genomic neighborhood around the RiPP BGCs in *C. michiganensis* NCPPB 382 to identify homologue genes in other *Clavibacter* species. Using antiSMASH's annotation of the RiPP BGCs as a reference, the amino acid sequences of non-biosynthetic genes were used as query for CORASON increasing the *cluster_radio* value to 30. After identifying homologous genes in the BGCs' genomic neighborhood, Easyfig (Sullivan et al., 2011) and BLASTn (Zhang et al., 2004) were used to compare the neighboordhood's sequences at the nucleotide level. CORASON's output was used as a reference to color homologous genes in Easyfig's output figure.

Phenotypic characterization in planta of* C. michiganensis *genetic diversity.

Disease development assays were performed at a greenhouse in July 2020, June 2021 and August 2021. *Clavibacter*-free tomato plantlets were provided by a commercial supplier (Plantanova, Mexico) with three to four true leaves, which

were inoculated with selected *C. michiganensis* strains from our collection. Three non *C. michiganesis* strains, namely, *C. californiensis* CFBP 8216, *C. phaseoli* CFBP 8217 (both acquired from the corresponding type strain collection) and *C. capsici* RA1B (isolated by us from wild tomato) were also included. Bacterial inoculation was made by scraping the surface (0.5 cm^2) of the stem with a needle below the first two true leaves and then puncturing slightly at the center of the scrapped area. Inoculum consisted of 5 µl of bacterial culture with a concentration of 1.5×10^9 CFU. Mock-inoculated control plants were generated using sterile water. Disease development was monitored in a weekly fashion for 6 weeks after inoculation. Disease progression classification and disease index calculation were performed as in Valenzuela et al. (2021) for each week to obtain a disease progress curve. The so-called Area Under Disease Progress Curve (AUDPC) was calculated using the *audpc* function from the agricolae package in R (Mendiburu, 2022).

Results

C. michiganensis *prevails in commercial tomato cultivars but not in wild relatives.*

To our knowledge, tomato bacterial canker has only been reported in commercial tomato cultivars, which led us to search for *C. michiganensis* existing in tomato wild relatives. If so, these wild plants could act as a natural reservoir for pathogens jumping to related host plants. Moreover, if they existed, these bacteria could provide us with valuable insights on the evolution of *C. michiganensis* to adapt and thrive in tomato cultivars. Hence, we attempted to isolate *Clavibacter* strains from wild tomato populations. We sampled 39 plants from two different populations and isolated 222 bacterial strains on semi selective CMM1 medium, which has been reliably used for isolation of *C. michiganensis* (Alvarez et al., 2005). Based on their 16S rRNA gene, most of the strains isolated correspond to the *Micrococcales* bacterial order (**Supp. Table S1**). Unexpectedly, only one isolate, termed RA1B, belonged to the *Clavibacter* genus, which came from a wild tomato plant with no noticeable disease symptoms. The 16S rRNA indicated that RA1B belongs to the *Clavibacter* genus, while ANI (Average Nucleotide Identity) analysis performed after genome sequencing showed that RA1B resembles *C. capsici*: 90.29% percentage identity when compared with *C. michiganensis* NCPPB 382 *vs.* 98.93% when compared with *C. capsici* PF008 (**Supp. Fig. 2, Supp. Table S2**).

Clavibacter *has undergone multiple hosts shifts unrelated to wild tomatoes*.

To determine the phylogenetic relationships among our strains and other previously reported *Clavibacter* strains, we assembled a genus DB including diverse *Clavibacter* genomes (**Supp. Table S3**). Then, we reconstructed a phylogenomic tree using 1231 single copy genes found to be conserved in the 69 genomes. Even though *C. michiganensis* is commonly associated with tomato, the fact that we managed to isolate a *C. capsici* strain from a wild tomato variety, prompted us to analyze the relationship between the bacterial phylogeny and the hosts from which they were isolated (**Fig. 1**). Clades and subclades were defined

based on the tree topology and ANI values (**Supp. Fig. 2**, **Supp. Table S2**), where monophyletic strains with >97% and >93% ANI similarity were considered part of the same subclade or clade, respectively. Thus, *Clavibacter* strains grouped into ten different subclades, plus eleven single-strain lineages (SSLs) that remain to be populated with further isolates and their genome sequences.

The phylogeny revealed that *Clavibacter* strains come from a more diverse and broader host diversity than expected. Moreover, the host-strain relationship seems to be independent of the evolutionary history of the strains, as a lack of relationship between the species, their pathogenicity, or the position of each species in the tree, was found. For instance, the *Clavibacter* species that are pathogenic in plants of the same family do not cluster together but are situated in distantly related clades: *C. nebraskensis*, pathogenic in maize (a *Poaceae* plant) and *C. insidiosus*, pathogenic in alfalfa (a *Fabaceae*) are clustered into sister clades. In contrast, *C. capsici, C. sepedonicus* and *C. michiganensis*, which infect *Solanaceae* plants, are placed at early and late divergent subclades. Moreover, *C. michiganensis* and *C. sepedonicus* both diverged after a clade of *Poaceae* isolates we termed the Pre-Solanaceae Jump (PSJ) clade, since their ancestor diverged from the lineage that precedes both *Solanaceae*-related species. SSLs with no reported pathogenic activity (1, 4, 5, 6, 7, 10) form monophyletic clades with the pathogenic strains. SSLs 4 and 5 cluster with *C. insidiosus* and *C. nebraskensis* subclades, SSLs 6 and 7 cluster with *C. phaseoli* (formerly known as *C. chilensis*, Arizala et al., 2022) while SSL10 cluster with SSL9 and the *C. tessellarius* sub-clade. The case of the so-called “Broad *C. michiganensis* clade” (from now onwards BCm clade) is particularly interesting since the *C. michiganensis* pathogenic subclade clusters with three non-pathogenic groups: SSL1 and the Pre-Phytonotic event 1 and 2 subclades (PP1 and PP2, **Fig. 1**), which were named as such since they descend from the lineage that diverged from the one that originated *C. michiganensis*.

Except from the *C. nebraskensis* subclade, all the subclades consist of strains obtained from at least two different host species. While there are cases where the host species are closely related, even belonging to the same plant

329 genus (e.g., *C. insidiosus* strains), there are several others where strains grouping
330 together come from distantly related hosts, such as in the *C. tessellarius, C.*
331 *phaseoli* and *C. michiganensis* subclades. Notably, several *Clavibacter* strains
332 isolated from *Poaceae* and *Solanaceae* plants (particularly from tomato) are
333 distributed throughout several clades. In some cases, strains isolated from both
334 plant families clustered within the same clade, such as in the so-called “Early
335 Divergent clade” (from now on ED clade), and in the *C. tessellarius, C. phaseoli*
336 and *C. michiganensis* subclades (**Fig. 1**). This pattern is particularly noticeable in
337 the subclade of *C. michiganensis* species, which includes several cases of host
338 shifts. This large clade, whose richness reflects the fact that we and others have
339 sampled it thoroughly, includes several strains obtained from other plants other
340 than tomato, with many *Poaceae* host plants examples but not the RA1B isolate
341 from wild tomato.
342
343

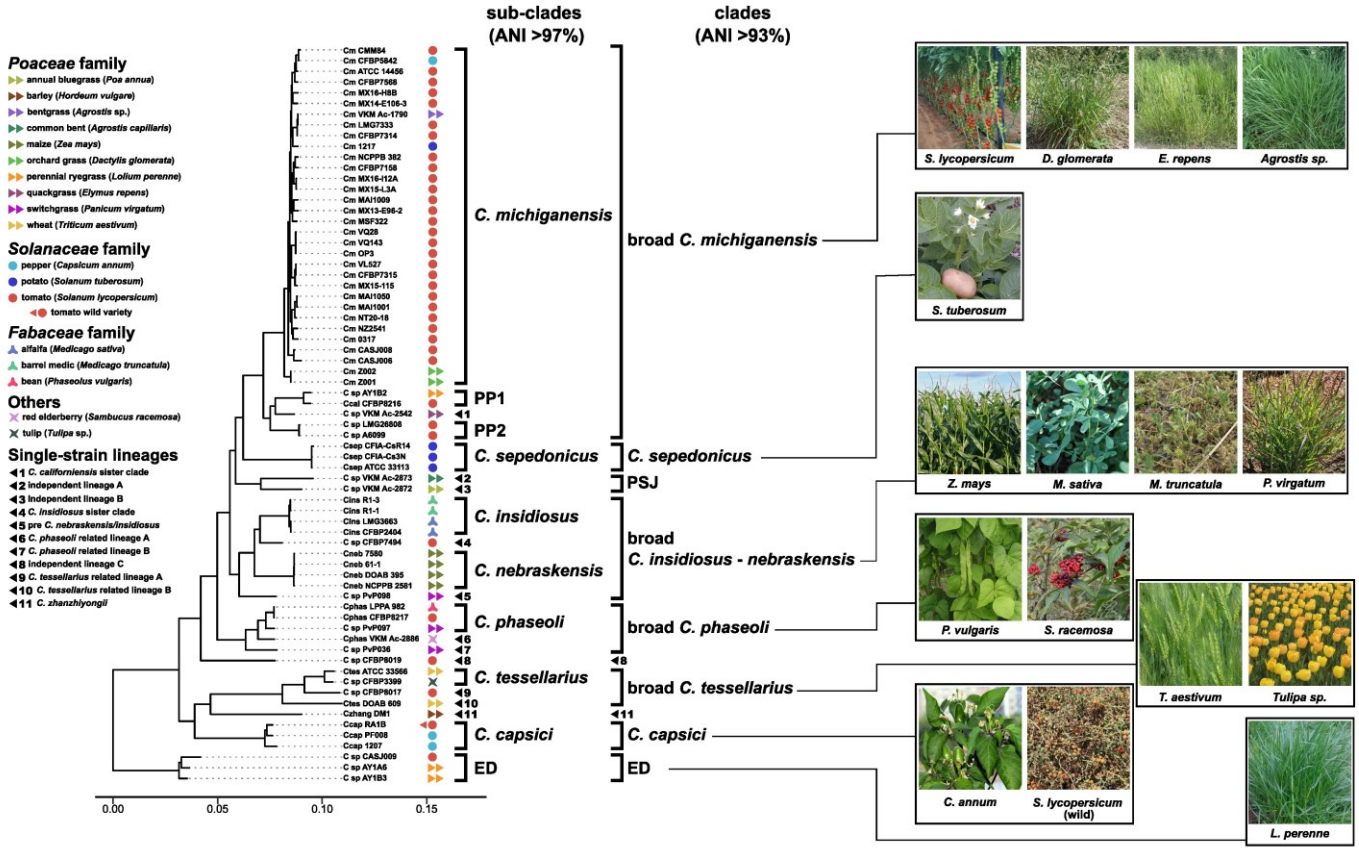

Figure 1. Phylogenomic tree of the *Clavibacter* genus and hosts of origin. Left: phylogenomic tree with clades and lineages indicated with brackets and numbers. Each host species is indicated by the colored geometrical shapes next to the strains' names. Right: representative pictures of plant hosts from which strains were isolated. Tree was inferred from 1231 core proteins. Only posterior probability support values <0.9 are shown. *Cm, C. michiganensis; C sp., Clavibacter* sp.; *Ccal, C. californiensis*; *Csep*, *C. sepedonicus*; *Cneb*, *C. nebraskensis*; *Cins, C. insidiosus*; *Cphas*, *C. phaseoli*; *Czhang, C. zhanzhiyongii*; *Ccap, C. capsici*; *Ctes, C. tessellarius*. ED = Early Divergent. PP1 = Pre-phytonotic event-1. PP2 = Pre-phytonotic event-2. PSJ = Pre-*Solanaceae* Jump.

351 ***Pathogenicity profiles of selected strains supports the phylogeny and host***
352 ***shifts.***

353 To test the phenotypic implications of our phylogeny, we then assessed the
354 pathogenicity of *C. capsici* RA1B in tomato plants and compare it to the two most
355 highly pathogenic strains of our *C. michiganensis* collection, strains MX15-115 and
356 MX16-I12A. Moreover, with these experiments, we also aimed at contrasting the
357 pathogenicity profiles of these known pathogens against the non michiganensis
358 *Clavibacter* strains isolated from tomato: *C. californiensis* CFBP 8216 and *C.*
359 *phaseoli* CFBP 8217, previously reported to be asymptomatic (Yasuhara-Bell &
360 Alvarez, 2015). These two strains, which provided a proxy to the grass strains (not
361 available to us), were referred to as the not-michiganensis tomato *Clavibacter*
362 (NMTC) and belong to the PP1 and *C. phaseoli* subclades (**Fig. 1**). These
363 *Clavibacter* strains were treated equally and used to infect tomato plants to
364 measure the development of disease symptoms, which were monitored for a
365 period of six weeks (**Fig. 2**). Throughout the period of evaluation, RA1B did not
366 show any sign of being pathogenic in tomato plants. Surprisingly, NMTC isolates
367 showed the development, even if very mild but yet quantifiable, of clear symptoms
368 compared to the *C. michiganensis* strains.
369

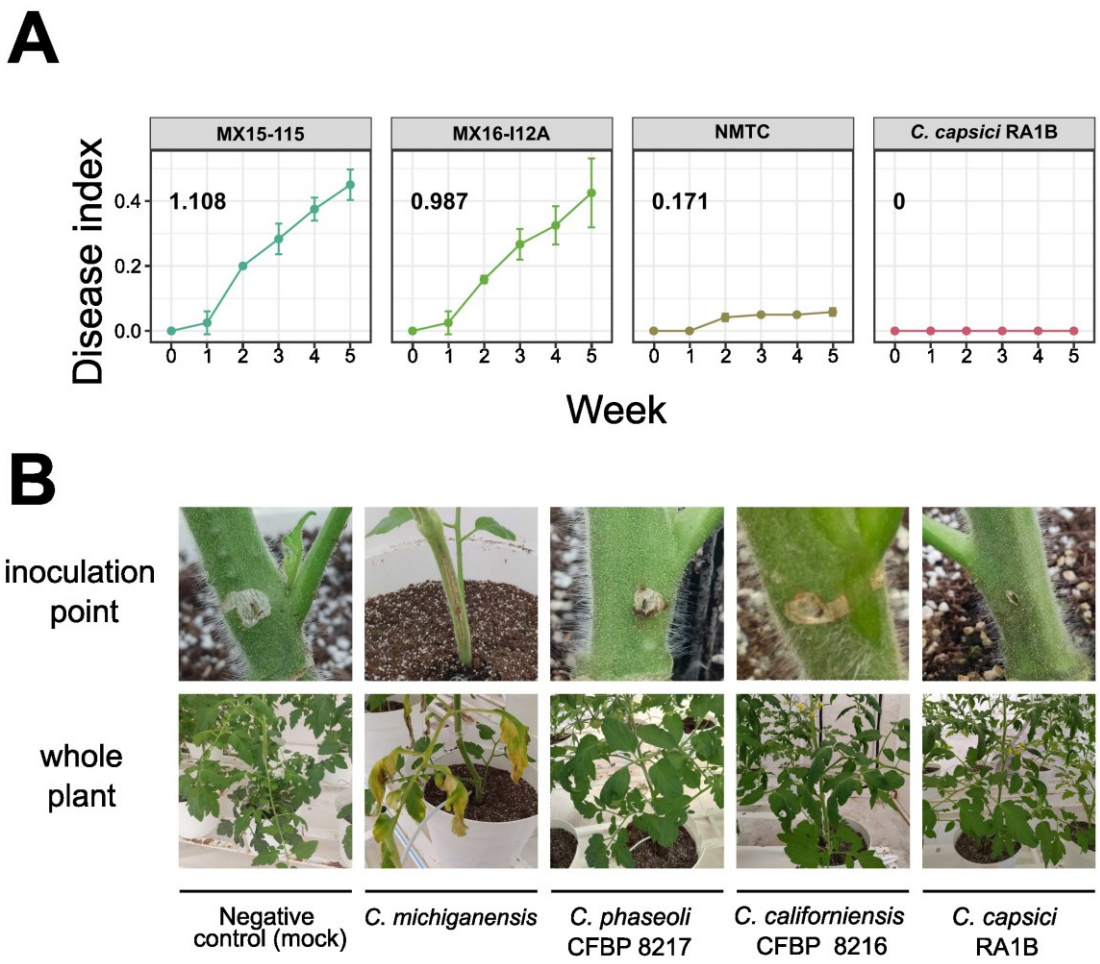

Figure 2. Pathogenicity profiling of selected *Clavibacter* strains. A. Pathogenicity profiling was quantitively done for selected strains using a disease index calculated for each treatment per week and used to calculate the area under disease progress curve (numbers in bold at top right of each subplot). **B.** Qualitative comparison of symptoms generated by selected *Clavibacter* strains, including *C. michiganensis* (strong symptoms), NMTC strains *C. phaseoli* and *C. californiensis* (mild symptoms only in the site of inoculation) and *C. capsici* RA1B (asymptomatic).

Genetic linkages between grass and commercial tomato isolates indicative of a host shift event.

When colonizing a new host, pre-adapted pathogens undergo changes that allow them to better thrive in their new environment. To identify these features at the genome level, we compared the gene content of key strains throughout the *Clavibacter* genus, and searched for genes that were associated with *C. michiganensis*. To achieve this, we performed a pangenomic analysis using the *Clavibacter* genus DB previously assembled. This analysis identified 11447 different gene families, of which 1766 were shared among all the strains in the dataset and 4183 were unique to one strain (singletons). Since our aim was to identify features common only to the *C. michiganensis* strains, conserved gene families and singletons were discarded, while the analysis continued with the 5498 remaining gene families. We then identified gene families acquired at the split of the *C. michiganensis* subclade and its sister clades within the BCm clade, as this would be the time of the occurrence of the proposed host shift. To test this hypothesis in more detailed, we employed two complementary approaches, as described next.

First, a gene enrichment analysis using Anvi'o was adopted to identify gene families enriched in the BCm clade and absent (or nearly absent) outside of it. This allowed us to filter out gene families shared by the strains belonging to the BCm clade from other clades of the *Clavibacter* genus. As most of the BCm clade consists of *C. michiganensis* strains, we expected that the gene families with the highest enrichment scores would be those with high degree of conservation in this species. Second, we identified gene families whose presence and absence pattern in the genus phylogeny had a phylogenetic signal. This allowed us discard gene families whose presence, although enriched in the BCm clade, were randomly distributed across it because of natural genomic variation. For this, we use the presence and absence of the gene families in the genomes as a binary trait whose phylogenetic signal could be tested under Fritz and Purvis' D. Conjunctly, these two approaches allowed us to identify 103 gene families (**Supp. Table 4**) that were

409 conserved in the BCm clade (**Fig. 3**). We further classified the conserved gene
410 families between those that were present outside the *C. michiganensis* clade from
411 those that were found exclusively in these species.

412 Operonic gene organization in bacteria is suggestive of functional
413 association between genes that cluster together on the same genomic region. For
414 this reason, we analyzed if the genes identified as conserved gene families co-
415 locate in the genome of *C. michiganensis*. Using strain NCPPB 382 as a reference,
416 we defined these genes' location, leading to the identification of nine regions or loci
417 (labeled from 1 to 8, and PAI) consisting of neighboring genes (**Fig. 3**, left). The
418 first and largest region identified corresponds to the pathogenicity island (PAI)
419 identified by Gartemann et al., (2008). Indeed, a total of 29 gene families
420 containing genes associated with the PAI were identified by our analysis. Only
421 three of the identified gene families belonging to the PAI had homologs outside the
422 *C. michiganensis* clade. One of them correspond to a family of hypothetical
423 proteins, while the other two correspond to gene families related to the AbiEii
424 toxin/antitoxin complex, which helps bacteria to deal with phages. Interestingly,
425 genes of these last two gene families can be found, besides the PAI, in the pCM2
426 plasmid (**Supp. Table 4**). Likewise, loci 5 was found to be the michiganin
427 biosynthetic gene cluster (BGC), known to encode for the synthesis of a
428 ribosomally synthesized and post-translationally modified peptide (RiPP) previously
429 isolated (Holtsmark et al., 2006) and used for PCR diagnostic purposes (Yasuhara-
430 Bell et al., 2014). These loci belong to the gene families with some occurrence
431 outside the *C. michiganensis* clade, albeit at a low frequency. Besides the PAI and
432 michiganin BGC, whose identification confirm the validity of the approach adopted,
433 we identified seven more loci indicative of the evolutionary dynamics leading to *C.*
434 *michiganensis*.

435 Three loci belonged to genes exclusively found in *C. michiganensis*, as the
436 PAI (i.e., 3, 4 and 8), and four loci included gene families present outside this
437 clade, like the michiganin BGC (i.e., 1, 2, 6 and 7). Functional annotation of the
438 genes conformed these novel *C. michiganensis* loci and predicted their putative
439 roles during evolution of this pathogen. These genes were found to classify in three

different types, (i) transport systems, (ii) RiPP BGCs, and (iii) undetermined (**Fig. 3**, right). Four out of the seven new loci identified (1, 2, 7 and 8) have genes suggestive of transport systems. Genes within loci 2 and 7 were annotated as glycosyl hydrolases, which suggests the transporters encoded in these loci participate in carbohydrate assimilation. Locus 1 was annotated as a Major Facilitator type transporter, but the substrates upon which this system may operate, could not be predicted. For annotation of the hypothetical or genes of unknown function included in locus 8, many of which were found to be conserved in other actinobacteria, we used ProteInfer. These analyses hinted towards nitrogen metabolism, including metabolites such as serine, as possible functions encoded by these genes. Also, an unprecedented RiPP BGC, *i.e.,* locus 4, could be annotated with antiSMASH. We termed the putative product of this BGC 'michivionin', as it shares the same chemical class (RiPP class III) with microvionin from *Microbacterium arborescens* 5913 (Wiebach et al., 2018), which is further analyzed in detail at the sequence level in the final section

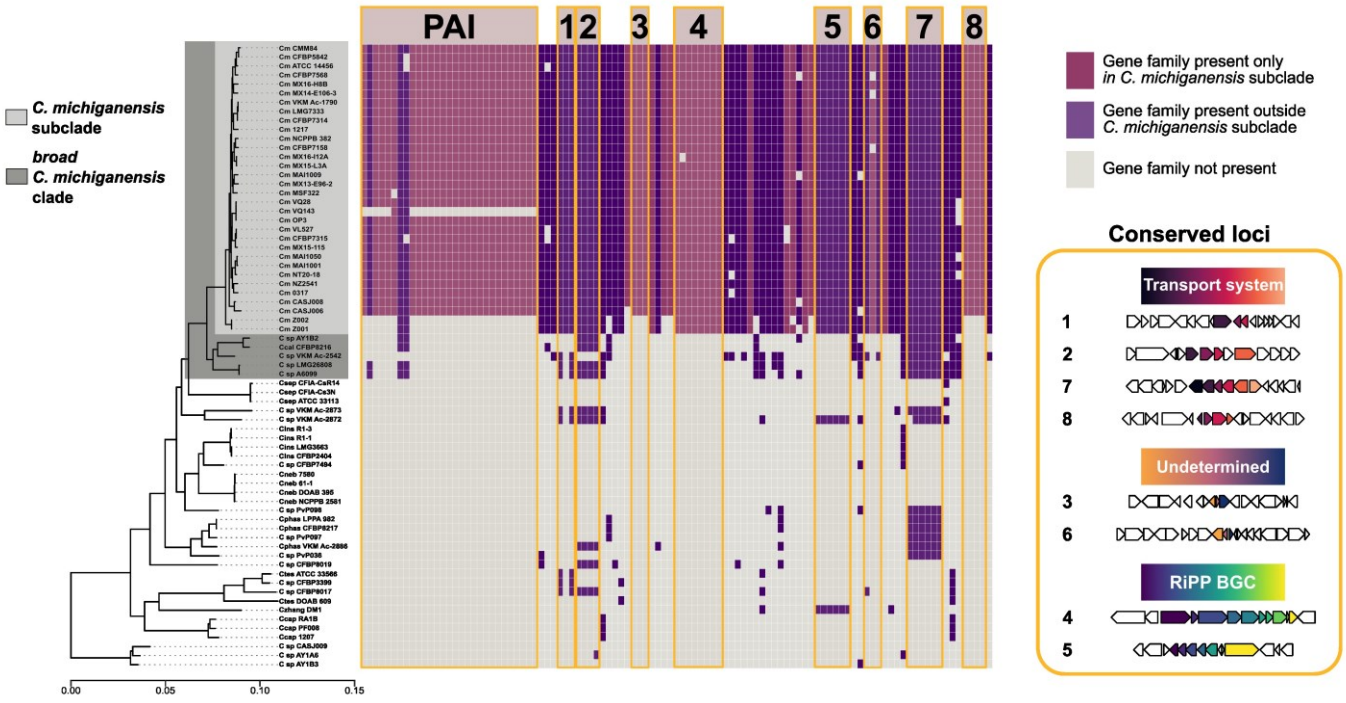

Figure 3. Conserved loci in *C. michiganensis* selected during the proposed host shift. Left, Heatmap of characteristic gene families of *C. michiganensis*. Columns of the heatmap represent different gene families in each of the analyzed strains throughout the phylogeny. The highlighted clade (dark gray) shows the BCm clade. Purple columns represent gene families found outside the *C. michiganensis* subclade, while magenta columns indicate gene families present exclusively in *C. michiganensis* subclade. Columns are ordered according to the occurrence of genes in the genome of *C. michiganensis* NCPPB 382. Gene families corresponding to the pathogenicity island are indicated as PAI. **Right.** Identified loci consisting of more than three contiguous, or semi contiguous genes, as described in Methods. Predicted functions of each gene in these loci is displayed below for each region. Associated function for each loci is indicated by a distinctive color gradient: purple → yellow, RiPPs; deep purpule → orange, transport systems; orange → blue, undetermined.

Genome mining of RiPPs implicated with* Clavibacter *adaptation to tomato.

Since the results of the pangenomic analysis suggested that RiPP BGCs may be a characteristic feature of *C. michiganensis*, we decided to explore the biosynthetic potential of the entire genus for these ecologically relevant metabolites (Li & Rebuffat, 2020). For this, we annotated the RiPP BGCs in the genus DB and group them into Gene Cluster Families (GCF) using BiG-SCAPE (**Fig. 4A**, top). We then resolved these GCFs throughout the genus phylogeny (**Fig. 4A,** bottom). Overall, we found twelve networks each one corresponding to a different GCF distributed across the whole genus. Only network I, corresponding to a Linear azol(in)e-containing peptide (LAP) RiPP BGC, was fully conserved throughout the genus. Along with this network, II, IV and IX were also found in two or more clades, *e.g.* network II in the *C. michiganensis* subclade, the SSL3 and in the *C. zhangzhiyongii* strain; network IV in the *C. sepedonicus* and *C. capsici* clades; and network IX in SSL1 and PCC2. Networks III, V and VI were found to be exclusive for one *Clavibacter* species each: *C. michiganensis*, *C. capsici* and *C. sepedonicus*, respectively, although the subclades of the latter two species are poorly populated, which is the case for the rest of the networks with only one strain. In contrast, the presence of networks II and III (**Fig. 4B**, top), which correspond to locus 5 (michiganin BGC) and locus 4 (michivionin BGC) from the previous pangenomic analysis, are what distinguishes *C. michiganensis* from the other *Clavibacter* species.

The previous analysis revealed that the michiganin BGC, once considered a distinctive feature of *C. michiganensis*, is shared with other *Clavibacter* species. Hence, we decided to inspect the genomic neighborhood of this BGC and compare it with other *Clavibacter* species to gain insights about the evolutionary history of this and other BGCs. Even though the michivionin BGC is exclusive to *C. michiganensis*, according to our results, we inspected and compared its genomic neighborhood as well. We selected thirteen *Clavibacter* strains representing main clades of the genus as resolved by the phylogenetic tree and performed a synteny analysis of the surrounding of both BGCs using CORASON (**Fig. 4B**, bottom).

Based on the identified homologous genes in each loci, we then used Easyfig and BLASTn to compare the nucleotide sequences of the different loci. BLASTn alignment for regions over 1000bp showed high degree of similarity for both neighborhoods: 89.68 ± 5.62% identity (e-value 3.18e-70 ± 6.13 e-69) for michiganin's BGC neighborhood and 90.60 ± 2.92% (e-value 0) for michivionin's BGC vicinity. Out of the two loci, the michivionin neighborhood is highly conserved in every *Clavibacter* strain. The sole difference was the presence of the michivionin BGC itself. In contrast, genomic neighborhood around the michiganin BGC is more variable, yet orthologous genes could still be identified throughout all strains. Interestingly, a michiganin-like BGC exists in *Microbacterium arborescens* 5913, the microvionin producer, coincidently in the vicinity of its cognate BGC. Comparison was then expanded to include this *Microbacterium* species (**Fig. 4B**), showing sequence and gene organization similarities only for the michiganin-related BGCs but not for the analogous microvionin and michivionin BGCs. Despite these, our analyses suggest a possible correlation between RiPPs BGCs, horizontally transferred amongst related plant *Micrococcales* genera, and evolution of *C. michiganensis* as a tomato pathogen from grass endophytes.

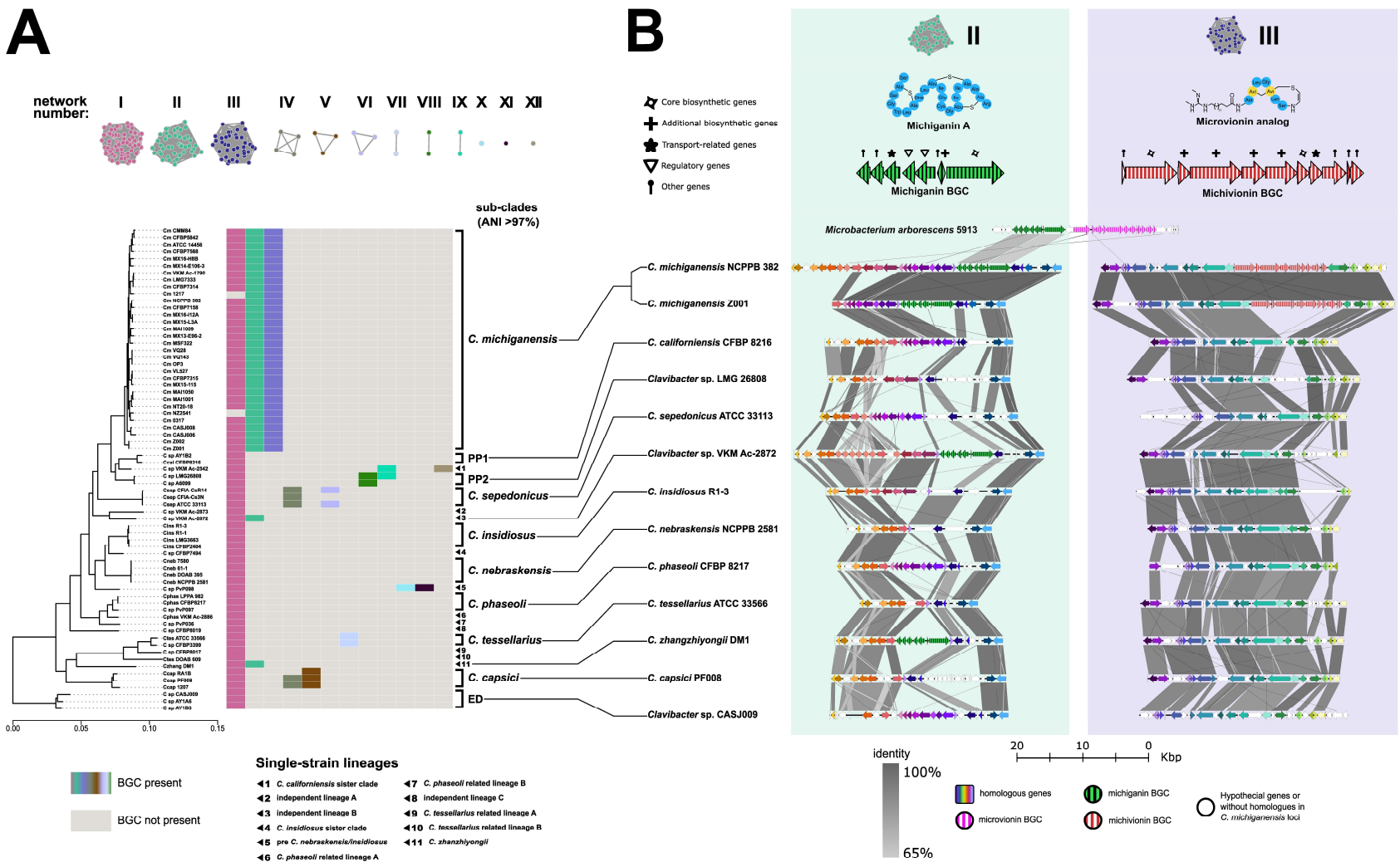
A
network number: I II III IV V VI VII VIII IX X XI XII
sub-clades (ANI >97%)
C. michiganensis
PP1
PP2
C. sepedonicus
C. insidiosus
C. nebraskensis
C. phaseoli
C. tessellarius
C. capsici
ED
C. michiganensis NCPPB 382
C. michiganensis Z001
C. californiensis CFBP 8216
Clavibacter sp. LMG 26808
C. sepedonicus ATCC 33113
Clavibacter sp. VKM Ac-2872
C. insidiosus R1-3
C. nebraskensis NCPPB 2581
C. phaseoli CFBP 8217
C. tessellarius ATCC 33566
C. zhangzhiyongii DM1
C. capsici PF008
Clavibacter sp. CASJ009
0.00 0.05 0.10 0.15
BGC present
BGC not present
Single-strain lineages
1 C. californiensis sister clade
2 independent lineage A
3 independent lineage B
4 C. insidiosus sister clade
5 pre C. nebraskensis/insidiosus
6 C. phaseoli related lineage A
7 C. phaseoli related lineage B
8 independent lineage C
9 C. tessellarius related lineage A
10 C. tessellarius related lineage B
11 C. zhanzhiyongii
B
Core biosynthetic genes
Additional biosynthetic genes
Transport-related genes
Regulatory genes
Other genes
II
Michiganin A
Michiganin BGC
III
Microvionin analog
Michivionin BGC
Microbacterium arborescens 5913
identity
100%
65%
20 10 0 Kbp
homologous genes
michiganin BGC
microvionin BGC
michivionin BGC
Hypothecial genes or without homologues in C. michiganensis loci

Figure 4. Predicted RiPP BGCs in *Clavibacter*. A. Top: RiPP BGC networks representing a different set of GCFs. Bottom: Presence of each BGC network in the *Clavibacter* genus, with each column displaying the presence or absence of each BGC as they appear in the top part of the panel. ED = Early Divergent. PP1 = Pre-Phytonotic event-1. PP2 = Pre-Phytoonotic event-2. **B.** Top: RiPP BGCs specific to the *C. michiganensis* subclade. The BGC from *C. michiganensis* NCPPB 382 is shown. Representative structures of the type of RiPP expected are shown in the right-hand side. Bottom: Genomic neighborhood of michiganin and microvionin BGCs highlighting horizontal gene transfers. The genomic neighborhood of michiganin and michivionin BGCs of *C. michiganensis* NCPPB 382 compared to the same loci in thirteen selected *Clavibacter* strains. Same colors are used for homologous genes believed to be orthologous based on sequence similarity and shared synteny. The microvionin BGC from *Microbacterium arborescens* 5913, which includes a michiganin-like BGC in its vicinity, is displayed on the top of the other loci.

Discussion

The *Clavibacter* genus has been known for a long time, mainly because of their pathogenic members. With some exceptions, it seems that the strains studied with genomics throughout the years have been isolated from the very same hosts these bacteria can cause disease in highly relevant crops. However, there are several reports of *Clavibacter* strains being isolated from a variety of plant sources like plum (Janisiewicz et al., 2013), coffee (Vega et al., 2005), poplar (Ulrich et al., 2008) and even wild plants (Ding et al., 2011; Zinniel et al., 2002). Moreover, several of the strains included in our analysis were isolated from sources different to those commonly associated with these pathogenic species. Hence, *Clavibacter* bacteria are not restricted to agricultural systems, and they can thrive in wild plant populations, as evidenced by the isolation of RA1B from a wild variety of tomato.

Unexpectedly, plants belonging to the *Poaceae* and *Solanaceae* plant families are the most frequent hosts of *Clavibacter* strains. These plant families are not closely related as shown by previous phylogenetic reconstructions (Leebens-Mack et al., 2019; Särkinen et al., 2013; Soreng et al., 2017), which clearly contrast with the close association between several *Clavibacter* strains obtained from such diverse hosts. The occurrence of such relationships throughout the phylogeny, at subclade and clade levels, and the incongruence between the *Clavibacter* genus phylogeny and those of their hosts, indicates lack of plant-bacteria co-evolution. Host shift is seen as a mechanism used by bacteria for long-term survival, as it allows pathogens to evolve and diversify through radiation and speciation (Langridge et al., 2015; Thines, 2019). Since host-shifting seems to be common trend in the Clavibacter genus, as shown herein, its very likely that the known pathogenic species appeared because of such mechanism after host shifting events. Even though *C. michiganensis* is the *Clavibacter* species typically associated with tomato (*S. lycopersicum*) the existence of several tomato isolates not belonging to this species and spread throughout the genus phylogeny is indicative of this phenomenon. As several conditions are required to perform a successful host shift, this observation suggests pre-adaptation as well as the occurrence of a series of events allowing these bacteria to overcome ecological

barriers and colonize tomato plants, which could be related to current production practices (Anzalone et al., 2022; Hao et al., 2019; Ristaino et al., 2021; Wyngaard & Kissinger, 2022) or the impact of domestication and breeding on the plant's physiology and its microbiome (Carrillo et al., 2019; Jaiswal et al., 2020; Soldan et al., 2021). The existence of *C. michiganensis* strains isolated from grasses at the base of this subclade (**Fig. 1**) supports the hypothesis of *C. michiganensis*' origin outside of tomato. This is further supported by our inability to isolate *C. michiganensis* from wild tomato plants, both in Mexico and in Chile (M. Valenzuela, unpublished).

Based on in-depth genomics analyses we were able to identify the mark of key signatures of adaptation in the genetic composition of *C. michiganensis*. Our pangenomic analysis highlights that the acquisition of genes represents a key turning point during the evolution of *C. michiganensis* as a tomato pathogen, providing experimentally testable hypotheses. For instance, mutants lacking the PAI, whose genes were identified by our analysis, have been shown to have reduced virulence (Gartemann et al., 2008). Interestingly, *C. michiganensis* strain VQ143, which lacks most of the PAI genes we identified, has shown a low degree of virulence when tested *in planta* (Valenzuela et al., 2021). Identification of conserved gene families in *C. michiganensis* encoding for carbohydrate and nitrogen compounds transporter proteins, suggests that nutrient acquisition strategies were relevant for the adaptation of this bacteria to the tomato xylem environment. Given that nutrient acquiring adaptations are key for bacterial pathogens, and endophytes alike, to thrive in the poor nutrient environment provided by the xylem, nutrients are not only used for metabolism but also as signals that can trigger environmental-driven specific responses (De La Fuente et al., 2022).

The occurrence of RiPP encoding BGCs as a distinctive feature of *C. michiganensis* and related species suggests that dealing with bacterial competitors was an important adaptation to the tomato environment. It has been reported that, as in the case of michiganin, the molecules produced by these BGCs have the capacity to inhibit the growth of closely related bacteria, including other *Clavibacter*

589 species (Holtsmark et al., 2006). However, the compounds produced by RiPP
590 BGCs could have other roles different than antibiosis, as their ability to mediate
591 intra-specific and host-bacteria interactions is well-acknowledged (Li & Rebuffat,
592 2020). The most parsimonious mechanism for spreading of these BGCs seems
593 horizontal gene transfer, again speaking out of the contribution of RiPPs towards
594 adaptation and evolution of pathogenic lifestyle.

595 In summary, our long-term and in-depth evolutionary genomics analyses –
596 including new data doubling the number of *C. michiganensis* genome sequences
597 available– solves a long-standing mystery about the origin and unpredictable
598 pathogenic behavior of *C. michiganensis* causing bacterial canker. A better
599 understanding of the evolutionary history of this seed-borne pathogen, explained
600 by a host shift from a grass to tomato, lightens up several possibilities for its control
601 and diagnostic, and for avoiding the occurrence of similar scenarios involving other
602 *Clavibacter* species (*e.g. C. nebraskensis*; Osdaghi et al., 2023; Webster et al.,
603 2019). On one hand, we anticipate that experimental validation of the candidate
604 genes we have identified here will provide a complete picture of the pathogenicity
605 of this endophyte. This in turn will make appearance of the disease in crops more
606 predictable, traceable, and ultimately, controllable. On the other hand, it also
607 provides lessons about pitfalls during plant genetic breeding, which represents the
608 most likely place and moment for a host shift involving a seed-borne endophyte to
609 occur. In this respect, fighting the enemy from within, *e.g.* by incorporating plant
610 rewilding (Chen et al., 2017) and/or microbiome engineering (Compant et al., 2019;
611 Noman et al., 2021) strategies, seems a more sustainable solution than treating
612 this phytopathogen as an opportunist with the high costs associated with a 'search
613 and destroy' strategy based on disinfectants, so-called certified seeds and/or
614 agrochemicals.
615
616

Data availability.

Genomes from *Clavibacter* isolates obtained in Mexico, Uruguay and the Netherlands used for this research will be released as part of the BioProject PRJNA996097.

Attributions

The following plant pictures in Figure 1 were obtained from these sources: *S. lycopersicum* from "Greenhouse Israel_IMG_3118.JPG" by Eddau, which is licensed under CC BY-SA 3.0; *D. glomerata* from "DactylisGlomerataIreland.JPG" by Notafly2, which is licensed under CC BY-SA 3.0; *E. repens* from "Elymus repens (3738612411).jpg" by Matt Lavin, which is licensed under CC BY-SA 2.0; *Agrostis sp.* from "Gewoon struisgras Agrostis tenuis.jpg" by Rasbak, which is licensed under CC BY-SA 3.0; *S. tuberosum* from "Starr 020701-0018 Solanum tuberosum.jpg" by Forest & Kim Starr, which is licensed under CC BY 3.0; *Z. mays* from "ZeaMays.jpg" by Christian Fischer, which is licensed under CC BY-SA 3.0; *M. sativa* from "Medicago sativa Alfalfa ლურჯი იონჯა.JPG" by Lazaregagnidze, which is licensed under CC BY-SA 4.0; *M. truncatula* from "Medicago truncatula habit1 Denman - 30067303043.jpg" by Macleay Grass Man, which is licensed under CC BY 2.0; *P. virgatum* from "Panicum virgatum Shenandoah 13zz.jpg" by David J. Stang, which is licensed under CC BY-SA 4.0**;** *P. vulgaris* from "Snijboon peulen Phaseolus vulgaris.jpg" by Rasbak; which is licensed under CC BY-SA 3.0; *S. racemosa* from "Red-berried Elder (Sambucus racemosa) - Oslo, Norway 2020-08-03 (02).jpg" by Ryan Hodnett; which is licensed under CC BY-SA 4.0; *T. aestivum* from "Starr-160415-0766-Triticum aestivum-home grown-Hawea Pl Olinda-Maui (26352044963).jpg" by Forest & Kim Starr, which is licensed under CC BY 3.0 US ; *Tulipa sp.* from "Tulipa Golden Apeldoorn.jpg" by elPadawan, which is licensed under CC BY-SA 2.0; *C. annum* from "Sibirischer Hauspaprika mit noch grünen Chilis.jpg" by Singlespeedfahrer, which is licensed under CC0 1.0 and *L. perenne* from "Lolium perenne Engels raaigras.jpg" by Rasbak, which is licensed under CC BY-SA 3.0. All images were solely cropped to the same size, no other modifications were performed.

Acknowledgements

This work was funded by Koppert Mexico SA de CV, Endogenomiks SAPI and Conacyt, Mexico (CB-2016 285746). FBG gratefully acknowledges support from Koppert BV, the Netherlands, during a sabbatical stay; and AYO from Conacyt, Mexico (748806) and the Institute of Biology, Leiden University, to complete PhD research. We are also grateful to Jose Luis Steffani Vallejo, Leslie Martinez Silva and Martha Angélica Martínez Tovar (Mexico) for their technical support in the greenhouse assays, and to Roberto Aviña Carlin and Alvaro Nieto for field work support.

References

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research*, *25*(17), 3389–3402. https://doi.org/10.1093/nar/25.17.3389

Alvarez, A. M., Kaneshiro, W. S., & Vine, B. G. (2005). Diversity of Clavibacter michiganensis subsp. michiganensis Populations in Tomato Seed : What is the Significance ? *Environmental Protection*, 205–214.

Anzalone, A., Mosca, A., Dimaria, G., Nicotra, D., Tessitori, M., Privitera, G. F., Pulvirenti, A., Leonardi, C., & Catara, V. (2022). Soil and Soilless Tomato Cultivation Promote Different Microbial Communities That Provide New Models for Future Crop Interventions. *International Journal of Molecular Sciences*, *23*(15). https://doi.org/10.3390/ijms23158820

Arizala, D., Dobhal, S., Alvarez, A. M., & Arif, M. (2022). Elevation of Clavibacter michiganensis subsp. californiensis to species level as Clavibacter californiensis sp. nov., merging and re-classification of Clavibacter michiganensis subsp. chilensis and Clavibacter michiganensis subsp. phaseoli as Clavibacter. *International Journal of Systematic and Evolutionary Microbiology*, *72*(9), 005427. https://doi.org/10.1099/ijsem.0.005427

Blin, K., Shaw, S., Kloosterman, A. M., Charlop-Powers, Z., Van Wezel, G. P., Medema, M. H., & Weber, T. (2021). antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Research*, *49*(W1), W29–W35. https://doi.org/10.1093/NAR/GKAB335

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., Olson, R., Overbeek, R., Parrello, B., Pusch, G. D., Shukla, M., Thomason, J. A., Stevens, R., Vonstein, V., Wattam, A. R., & Xia, F. (2015). RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Scientific Reports 2015 5:1*, *5*(1), 1–6. https://doi.org/10.1038/srep08365

Carrillo, J., Ingwell, L. L., Li, X., & Kaplan, I. (2019). Domesticated tomatoes are more vulnerable to negative plant–soil feedbacks than their wild relatives. *Journal of Ecology*, *107*(4), 1753–1766. https://doi.org/10.1111/1365-2745.13157

Chen, Y. H., Shapiro, L. R., Benrey, B., & Cibrián-Jaramillo, A. (2017). Back to the origin: In situ studies are needed to understand selection during crop diversification. *Frontiers in Ecology and Evolution*, *5*(OCT), 1–8. https://doi.org/10.3389/fevo.2017.00125

Compant, S., Samad, A., Faist, H., & Sessitsch, A. (2019). A review on the plant microbiome: Ecology, functions, and emerging trends in microbial application. In *Journal of Advanced Research* (Vol. 19, pp. 29–37). Elsevier B.V. https://doi.org/10.1016/j.jare.2019.03.004

Davis, J. J., Wattam, A. R., Aziz, R. K., Brettin, T., Butler, R., Butler, R. M., Chlenski, P., Conrad, N., Dickerman, A., Dietrich, E. M., Gabbard, J. L., Gerdes, S., Guard, A., Kenyon, R. W., MacHi, D., Mao, C., Murphy-Olson, D., Nguyen, M., Nordberg, E. K., … Stevens, R. (2020). The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities. *Nucleic Acids Research*, *48*(D1), D606–D612. https://doi.org/10.1093/NAR/GKZ943

De La Fuente, L., Merfa, M. V., Cobine, P. A., & Coleman, J. J. (2022). Pathogen Adaptation to the Xylem Environment. *Annual Review of Phytopathology*, *60*, 163–186. https://doi.org/10.1146/annurev-phyto-021021-041716

Ding, S., Huang, C.-L., Sheng, H.-M., Song, C.-L., Li, Y.-B., & An, L.-Z. (2011). Effect of inoculation

with the endophyte Clavibacter sp. strain Enf12 on chilling tolerance in Chorispora bungeana. *Physiologia Plantarum*, *141*(2), 141–151. https://doi.org/10.1111/j.1399-3054.2010.01428.x

Dreier, J., Meletzus, D., & Eichenlaub, R. (1997). Characterization of the Plasmid Encoded Virulence Region pat -1 of Phytopathogenic Clavibacter michiganensis subsp. michiganensis. *Molecular Plant-Microbe Interactions®*, *10*(2), 195–206. https://doi.org/10.1094/MPMI.1997.10.2.195

Eichenlaub, R., Gartemann, K.-H., & Burger, A. (2006). Clavibacter michiganensis, a group of gram-positive phytopathogenic bacteria. In *Plant-Associated Bacteria* (pp. 385–421). Springer Netherlands. https://doi.org/10.1007/978-1-4020-4538-7_12

EPPO. (2016). PM 7/42 (3) Clavibacter michiganensis subsp. michiganensis. *EPPO Bulletin*, *46*(2), 202–225. https://doi.org/10.1111/epp.12302

Eren, A. M., Kiefl, E., Shaiber, A., Veseli, I., Miller, S. E., Schechter, M. S., Fink, I., Pan, J. N., Yousef, M., Fogarty, E. C., Trigodet, F., Watson, A. R., Esen, Ö. C., Moore, R. M., Clayssen, Q., Lee, M. D., Kivenson, V., Graham, E. D., Merrill, B. D., … Willis, A. D. (2021). Community-led, integrated, reproducible multi-omics with anvi'o. *Nature Microbiology*, *6*(1), 3–6. https://doi.org/10.1038/s41564-020-00834-3

Fritz, S. A., & Purvis, A. (2010). Selectivity in Mammalian Extinction Risk and Threat Types: a New Measure of Phylogenetic Signal Strength in Binary Traits. *Conservation Biology*, *24*(4), 1042–1051. https://doi.org/10.1111/j.1523-1739.2010.01455.x

Gartemann, K.-H., Abt, B., Bekel, T., Burger, A., Engemann, J., Flügel, M., Gaigalat, L., Goesmann, A., Gräfen, I., Kalinowski, J., Kaup, O., Kirchner, O., Krause, L., Linke, B., McHardy, A., Meyer, F., Pohle, S., Rückert, C., Schneiker, S., … Bartels, D. (2008). The Genome Sequence of the Tomato-Pathogenic Actinomycete Clavibacter michiganensis subsp. michiganensis NCPPB382 Reveals a Large Island Involved in Pathogenicity. *Journal of Bacteriology*, *190*(6), 2138–2149. https://doi.org/10.1128/JB.01595-07

Gitaitis, R. D. (1991). Detection of Clavibacter michiganensis subsp. michiganensis in Symptomless Tomato Transplants. *Plant Disease*, *75*(8), 834. https://doi.org/10.1094/PD-75-0834

Hao, Z., Chen, B., & Li, X. (2019). Relationship between soil chemical properties and microbial metabolic patterns in intensive greenhouse tomato production systems. *Https://Doi.Org/10.1080/03650340.2019.1666369*, *66*(10), 1334–1343. https://doi.org/10.1080/03650340.2019.1666369

Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*, *35*(2), 518–522. https://doi.org/10.1093/MOLBEV/MSX281

Holtsmark, I., Mantzilas, D., Eijsink, V. G. H., & Brurberg, M. B. (2006). Purification, characterization, and gene sequence of michiganin A, an actagardine-like lantibiotic produced by the tomato pathogen Clavibacter michiganensis subsp. michiganensis. *Applied and Environmental Microbiology*, *72*(9), 5814–5821. https://doi.org/10.1128/AEM.00639-06

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. https://doi.org/10.1094/MPMI.2000.13.7.703

Jaiswal, A. K., Mengiste, T. D., Myers, J. R., Egel, D. S., & Hoagland, L. A. (2020). Tomato Domestication Attenuated Responsiveness to a Beneficial Soil Microbe for Plant Growth Promotion and Induction of Systemic Resistance to Foliar Pathogens. *Frontiers in Microbiology*, *11*, 3309. https://doi.org/10.3389/fmicb.2020.604566

Janisiewicz, W. J., Jurick, W. M., Vico, I., Peter, K. A., & Buyer, J. S. (2013). Culturable bacteria

752 from plum fruit surfaces and their potential for controlling brown rot after harvest. *Postharvest*
753 *Biology and Technology*, *76*, 145–151. https://doi.org/10.1016/j.postharvbio.2012.10.004

754 Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Von Haeseler, A., & Jermiin, L. S. (2017).
755 ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*,
756 *14*(6), 587–589. https://doi.org/10.1038/nmeth.4285

757 Kleitman, F., Barash, I., Burger, A., Iraki, N., Falah, Y., Sessa, G., Weinthal, D., Chalupowicz, L.,
758 Gartemann, K.-H., Eichenlaub, R., & Manulis-Sasson, S. (2008). Characterization of a
759 Clavibacter michiganensis subsp. michiganensis population in Israel. *European Journal of*
760 *Plant Pathology*, *121*(4), 463–475. https://doi.org/10.1007/s10658-007-9264-z

761 Kloosterman, A. M., Shelton, K. E., van Wezel, G. P., Medema, M. H., & Mitchell, D. A. (2020).
762 RRE-Finder: a Genome-Mining Tool for Class-Independent RiPP Discovery. *MSystems*, *5*(5).
763 https://doi.org/10.1128/MSYSTEMS.00267-20/SUPPL_FILE/MSYSTEMS.00267-20-
764 SF007.PDF

765 Langridge, G. C., Fookesa, M., Connora, T. R., Feltwella, T., Feaseya, N., Parsonsb, B. N., Seth-
766 Smitha, H. M. B., Barquista, L., Stedmana, A., Humphreyd, T., Wigleyb, P., Petersf, S. E.,
767 Maskellf, D. J., Coranderg, J., Chabalgoityh, J. A., Barrowi, P., Parkhilla, J., Dougana, G., &
768 Thomsona, N. R. (2015). Patterns of genome evolution that have accompanied host
769 adaptation in Salmonella. *Proceedings of the National Academy of Sciences of the United*
770 *States of America*, *112*(3), 863–868. https://doi.org/10.1073/pnas.1416707112

771 Le, S. Q., & Gascuel, O. (2008). An Improved General Amino Acid Replacement Matrix. *Molecular*
772 *Biology and Evolution*, *25*(7), 1307–1320. https://doi.org/10.1093/MOLBEV/MSN067

773 Leebens-Mack, J. H., Barker, M. S., Carpenter, E. J., Deyholos, M. K., Gitzendanner, M. A.,
774 Graham, S. W., Grosse, I., Li, Z., Melkonian, M., Mirarab, S., Porsch, M., Quint, M., Rensing,
775 S. A., Soltis, D. E., Soltis, P. S., Stevenson, D. W., Ullrich, K. K., Wickett, N. J., DeGironimo,
776 L., … Ka-Shu Wong, G. (2019). One thousand plant transcriptomes and the phylogenomics of
777 green plants. *Nature*, *574*(7780), 679–685. https://doi.org/10.1038/s41586-019-1693-2

778 Li, X., Tambong, J., Yuan, K. X., Chen, W., Xu, H., Lévesque, C. A., & De Boer, S. H. (2018). Re-
779 classification of Clavibacter michiganensis subspecies on the basis of whole-genome and
780 multi-locus sequence analyses. *International Journal of Systematic and Evolutionary*
781 *Microbiology*, *68*(1), 234–240. https://doi.org/10.1099/ijsem.0.002492

782 Li, Y., & Rebuffat, S. (2020). The manifold roles of microbial ribosomal peptide-based natural
783 products in physiology and ecology. *Journal of Biological Chemistry*, *295*(1), 34–54.
784 https://doi.org/10.1074/jbc.REV119.006545

785 Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., Dow, M., Verdier, V.,
786 Beer, S. V., Machado, M. A., Toth, I., Salmond, G., & Foster, G. D. (2012). Top 10 plant
787 pathogenic bacteria in molecular plant pathology. *Molecular Plant Pathology*, *13*(6), 614–629.
788 https://doi.org/10.1111/j.1364-3703.2012.00804.x

789 McCann, H. C. (2020). Skirmish or war: the emergence of agricultural plant pathogens. *Current*
790 *Opinion in Plant Biology*, *56*, 147–152. https://doi.org/10.1016/j.pbi.2020.06.003

791 McDonald, B. A., & Stukenbrock, E. H. (2016). Rapid emergence of pathogens in agro-ecosystems:
792 global threats to agricultural sustainability and food security. *Philosophical Transactions of the*
793 *Royal Society B: Biological Sciences*, *371*(1709), 20160026.
794 https://doi.org/10.1098/rstb.2016.0026

795 Meletzus, D., Bermphol, A., Dreier, J., & Eichenlaub, R. (1993). Evidence for plasmid-encoded
796 virulence factors in the phytopathogenic bacterium Clavibacter michiganensis subsp.
797 michiganensis NCPPB382. *Journal of Bacteriology*, *175*(7), 2131–2136.
798 https://doi.org/10.1128/jb.175.7.2131-2136.1993

799 Mendiburu, F. (2022). Agricolae: Statistical Procedures for Agricultural Research. *Statistical*
800 *Procedures for Agricultural Research*.

801 Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A.,
802 Lanfear, R., & Teeling, E. (2020). IQ-TREE 2: New Models and Efficient Methods for
803 Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, *37*(5), 1530–
804 1534. https://doi.org/10.1093/MOLBEV/MSAA015

805 Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L.,
806 Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., Finn, R. D., & Bateman, A. (2021).
807 Pfam: The protein families database in 2021. *Nucleic Acids Research*, *49*(D1), D412–D419.
808 https://doi.org/10.1093/NAR/GKAA913

809 Navarro-Muñoz, J. C., Selem-Mojica, N., Mullowney, M. W., Kautsar, S. A., Tryon, J. H., Parkinson,
810 E. I., De Los Santos, E. L. C., Yeong, M., Cruz-Morales, P., Abubucker, S., Roeters, A.,
811 Lokhorst, W., Fernandez-Guerra, A., Cappelini, L. T. D., Goering, A. W., Thomson, R. J.,
812 Metcalf, W. W., Kelleher, N. L., Barona-Gomez, F., & Medema, M. H. (2020). A computational
813 framework to explore large-scale biosynthetic diversity. *Nature Chemical Biology*, *16*(1), 60–
814 68. https://doi.org/10.1038/s41589-019-0400-9

815 Noman, M., Ahmed, T., Ijaz, U., Shahid, M., Azizullah, Li, D., Manzoor, I., & Song, F. (2021). Plant–
816 Microbiome Crosstalk: Dawning from Composition and Assembly of Microbial Community to
817 Improvement of Disease Resilience in Plants. *International Journal of Molecular Sciences*,
818 *22*(13), 6852. https://doi.org/10.3390/ijms22136852

819 Oh, E.-J., Hwang, I. S., Park, I. W., & Oh, C.-S. (2022). Comparative Genome Analyses of
820 Clavibacter michiganensis Type Strain LMG7333T Reveal Distinct Gene Contents in Plasmids
821 From Other Clavibacter Species. *Frontiers in Microbiology*, *12*(February), 1–12.
822 https://doi.org/10.3389/fmicb.2021.793345

823 Orme, D., Freckleton, R., Thomas, G., Petzoldt, T., Fritz, S., Isaac, N., & Pearse, W. (2013).
824 CAPER: comparative analyses of phylogenetics and evolution in R. *Methods in Ecology and*
825 *Evolution*, *3*, 145–151.

826 Osdaghi, E., Rahimi, T., Taghavi, S. M., Ansari, M., Zarei, S., Portier, P., Briand, M., & Jacques, M.-
827 A. (2020). Comparative Genomics and Phylogenetic Analyses Suggest Several Novel Species
828 within the Genus Clavibacter , Including Nonpathogenic Tomato-Associated Strains. *Applied*
829 *and Environmental Microbiology*, *86*(6), 1–20. https://doi.org/10.1128/AEM.02873-19

830 Osdaghi, E., Robertson, A. E., Jackson-Ziems, T. A., Abachi, H., Li, X., & Harveson, R. M. (2023).
831 Clavibacter nebraskensis causing Goss's wilt of maize: Five decades of detaining the enemy
832 in the New World. *Molecular Plant Pathology*, *24*(7), 675–692.
833 https://doi.org/10.1111/MPP.13268

834 Poysa, V. (1993). Evaluation of tomato breeding lines resistant to bacterial canker. *Canadian*
835 *Journal of Plant Pathology*, *15*(4), 301–304. https://doi.org/10.1080/07060669309501927

836 Pritchard, L., Glover, R. H., Humphris, S., Elphinstone, J. G., & Toth, I. K. (2015). Genomics and
837 taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens.
838 *Analytical Methods*, *8*(1), 12–24. https://doi.org/10.1039/C5AY02550H

839 Ristaino, J. B., Anderson, P. K., Bebber, D. P., Brauman, K. A., Cunniffe, N. J., Fedoroff, N. V.,
840 Finegold, C., Garrett, K. A., Gilligan, C. A., Jones, C. M., Martin, M. D., MacDonald, G. K.,
841 Neenan, P., Records, A., Schmale, D. G., Tateosian, L., & Wei, Q. (2021). The persistent
842 threat of emerging plant disease pandemics to global food security. *Proceedings of the*
843 *National Academy of Sciences of the United States of America*, *118*(23), e2022239118.
844 https://doi.org/10.1073/PNAS.2022239118/SUPPL_FILE/PNAS.2022239118.SAPP.PDF

845 Sabin, N. S., Calliope, A. S., Simpson, S. V., Arima, H., Ito, H., Nishimura, T., & Yamamoto, T.
846 (2020). Implications of human activities for (re)emerging infectious diseases, including COVID-

19. *Journal of Physiological Anthropology*, *39*(1), 29. https://doi.org/10.1186/s40101-020-00239-5

Särkinen, T., Bohs, L., Olmstead, R. G., & Knapp, S. (2013). A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): a dated 1000-tip tree. *BMC Evolutionary Biology*, *13*(1), 214. https://doi.org/10.1186/1471-2148-13-214

Sen, Y., Van Der Wolf, J., Visser, R. G. F., & Van Heusden, S. (2015). Bacterial canker of tomato: Current knowledge of detection, Management, Resistance, And interactions. *Plant Disease*, *99*(1), 4–13. https://doi.org/10.1094/PDIS-05-14-0499-FE

Shaiber, A., Willis, A. D., Delmont, T. O., Roux, S., Chen, L. X., Schmid, A. C., Yousef, M., Watson, A. R., Lolans, K., Esen, Ö. C., Lee, S. T. M., Downey, N., Morrison, H. G., Dewhirst, F. E., Mark Welch, J. L., & Eren, A. M. (2020). Functional and genetic markers of niche partitioning among enigmatic members of the human oral microbiome. *Genome Biology*, *21*(1), 1–35. https://doi.org/10.1186/S13059-020-02195-W/METRICS

Sharabani, G., Manulis-Sasson, S., Chalupowicz, L., Borenstein, M., Shulhani, R., Lofthouse, M., Sofer, M., Frenkel, O., Dror, O., & Shtienberg, D. (2014). Temperature at the early stages of Clavibacter michiganensis subsp. michiganensis infection affects bacterial canker development and virulence gene expression. *Plant Pathology*, *63*(5), 1119–1129. https://doi.org/10.1111/ppa.12199

Sharabani, G., Shtienberg, D., Borenstein, M., Shulhani, R., Lofthouse, M., Sofer, M., Chalupowicz, L., Barel, V., & Manulis-Sasson, S. (2013). Effects of plant age on disease development and virulence of Clavibacter michiganensis subsp. michiganensis on tomato. *Plant Pathology*, *62*(5), 1114–1122. https://doi.org/10.1111/ppa.12013

Smith, E. F. (1910). A new tomato disease of economic importance. *Science*, *31*(803), 794–796.

Soldan, R., Fusi, M., Cardinale, M., Daffonchio, D., & Preston, G. M. (2021). The effect of plant domestication on host control of the microbiota. *Communications Biology*, *4*(1), 1–9. https://doi.org/10.1038/s42003-021-02467-6

Soreng, R. J., Peterson, P. M., Romaschenko, K., Davidse, G., Teisher, J. K., Clark, L. G., Barberá, P., Gillespie, L. J., & Zuloaga, F. O. (2017). A worldwide phylogenetic classification of the Poaceae (Gramineae) II: An update and a comparison of two 2015 classifications. *Journal of Systematics and Evolution*, *55*(4), 259–290. https://doi.org/10.1111/jse.12262

Stork, I., Gartemann, K. H., Burger, A., & Eichenlaub, R. (2008). A family of serine proteases of Clavibacter michiganensis subsp. michiganensis: chpC plays a role in colonization of the host plant tomato. *Molecular Plant Pathology*, *9*(5), 599–608. https://doi.org/10.1111/j.1364-3703.2008.00484.x

Stukenbrock, E. H., & McDonald, B. A. (2008). The Origins of Plant Pathogens in Agro-Ecosystems. *Annual Review of Phytopathology*, *46*(1), 75–100. https://doi.org/10.1146/annurev.phyto.010708.154114

Sullivan, M. J., Petty, N. K., & Beatson, S. A. (2011). Easyfig: a genome comparison visualizer. *Bioinformatics*, *27*(7), 1009–1010. https://doi.org/10.1093/BIOINFORMATICS/BTR039

Tambong, J. T. (2017). Comparative genomics of Clavibacter michiganensis subspecies, pathogens of important agricultural crops. *PLOS ONE*, *12*(3), e0172295. https://doi.org/10.1371/journal.pone.0172295

Tancos, M. A., Lange, H. W., & Smart, C. D. (2015). Characterizing the Genetic Diversity of the Clavibacter michiganensis subsp. michiganensis Population in New York. *Phytopathology®*, *105*(2), 169–179. https://doi.org/10.1094/PHYTO-06-14-0178-R

Terlouw, B. R., Blin, K., Navarro-Muñoz, J. C., Avalon, N. E., Chevrette, M. G., Egbert, S., Lee, S.,

Meijer, D., Recchia, M. J. J., Reitz, Z. L., van Santen, J. A., Selem-Mojica, N., Tørring, T., Zaroubi, L., Alanjary, M., Aleti, G., Aguilar, C., Al-Salihi, S. A. A., Augustijn, H. E., … Medema, M. H. (2022). MIBiG 3.0: a community-driven effort to annotate experimentally validated biosynthetic gene clusters. *Nucleic Acids Research*, epub ahead of print. https://doi.org/10.1093/nar/gkac1049

Thapa, S. P., O'Leary, M., Jacques, M.-A. A., Gilbertson, R. L., Coaker, G., O'Leary, M., Jacques, M.-A. A., Gilbertson, R. L., Coaker, G., O'Leary, M., Jacques, M.-A. A., Gilbertson, R. L., & Coaker, G. (2020). Comparative genomics to develop a specific multiplex PCR assay for detection of clavibacter michiganensis. *Phytopathology*, *110*(3), 556–566. https://doi.org/10.1094/PHYTO-10-19-0405-R

Thapa, S. P., Pattathil, S., Hahn, M. G., Jacques, M.-A., Gilbertson, R. L., & Coaker, G. (2017). Genomic Analysis of Clavibacter michiganensis Reveals Insight Into Virulence Strategies and Genetic Diversity of a Gram-Positive Bacterial Pathogen. *Molecular Plant-Microbe Interactions®*, *30*(10), 786–802. https://doi.org/10.1094/MPMI-06-17-0146-R

Thines, M. (2019). An evolutionary framework for host shifts – jumping ships for survival. *New Phytologist*, *224*(2), 605–617. https://doi.org/10.1111/nph.16092

Ulrich, K., Ulrich, A., & Ewald, D. (2008). Diversity of endophytic bacterial communities in poplar grown under field conditions. *FEMS Microbiology Ecology*, *63*(2), 169–180. https://doi.org/10.1111/j.1574-6941.2007.00419.x

Valenzuela, M., González, M., Velásquez, A., Dorta, F., Montenegro, I., Besoain, X., Salvà-Serra, F., Jaén-Luchoro, D., Moore, E. R. B., & Seeger, M. (2021). Analyses of virulence genes of clavibacter michiganensis subsp. Michiganensis strains reveal heterogeneity and deletions that correlate with pathogenicity. *Microorganisms*, *9*(7), 1530. https://doi.org/10.3390/microorganisms9071530

Van Dongen, S., & Abreu-Goodger, C. (2012). Using MCL to extract clusters from networks. *Methods in Molecular Biology (Clifton, N.J.)*, *804*, 281–295. https://doi.org/10.1007/978-1-61779-361-5_15

Vega, F. E., Pava-Ripoll, M., Posada, F., & Buyer, J. S. (2005). Endophytic bacteria inCoffea arabica L. *Journal of Basic Microbiology*, *45*(5), 371–380. https://doi.org/10.1002/jobm.200410551

Wassermann, E., Montecchia, M. S., Garaventa, V. S., Correa, O. S., & Romero, A. M. (2020). Virulence and pCM1 plasmid carriage are related to BOX-PCR fingerprint type in strains of Clavibacter michiganensis subsp. michiganensis that cause bacterial wilt and canker of tomato in Argentina. *Plant Pathology*, *69*(4), 723–732. https://doi.org/10.1111/ppa.13163

Webster, B. T., Curland, R. D., McNally, R. R., Ishimaru, C. A., & Malvick, D. K. (2019). Infection, survival, and growth of clavibacter nebraskensis on crop, weed, and prairie plant species. *Plant Disease*, *103*(8), 2108–2112. https://doi.org/10.1094/PDIS-02-19-0346-RE

Weisberg, A. J., Grünwald, N. J., Savory, E. A., Putnam, M. L., & Chang, J. H. (2021). Genomic Approaches to Plant-Pathogen Epidemiology and Diagnostics. *Annual Review of Phytopathology*, *59*(1), 311–332. https://doi.org/10.1146/annurev-phyto-020620-121736

Whelan, S., & Goldman, N. (2001). A General Empirical Model of Protein Evolution Derived from Multiple Protein Families Using a Maximum-Likelihood Approach. *Molecular Biology and Evolution*, *18*(5), 691–699. https://doi.org/10.1093/OXFORDJOURNALS.MOLBEV.A003851

Wiebach, V., Mainz, A., Siegert, M. A. J., Jungmann, N. A., Lesquame, G., Tirat, S., Dreux-Zigha, A., Aszodi, J., Le Beller, D., & Süssmuth, R. D. (2018). The anti-staphylococcal lipolanthines are ribosomally synthesized lipopeptides brief-communication. *Nature Chemical Biology*, *14*(7), 652–654. https://doi.org/10.1038/s41589-018-0068-6

940 Wyngaard, S. R., & Kissinger, M. (2022). Tomatoes from the desert: Environmental footprints and
941 sustainability potential in a changing world. *Frontiers in Sustainable Food Systems*, *6*.
942 https://doi.org/10.3389/fsufs.2022.994920

943 Yasuhara-Bell, J., & Alvarez, A. M. (2015). Seed-associated subspecies of the genus Clavibacter
944 are clearly distinguishable from Clavibacter michiganensis subsp. michiganensis. *International*
945 *Journal of Systematic and Evolutionary Microbiology*, *65*(Pt_3), 811–826.
946 https://doi.org/10.1099/ijs.0.000022

947 Yasuhara-Bell, J., Marrero, G., & Alvarez, A. M. (2014). Genes clvA, clvF and clvG are unique to
948 Clavibacter michiganensis subsp. michiganensis and highly conserved. *European Journal of*
949 *Plant Pathology*, *140*(4), 655–664. https://doi.org/10.1007/s10658-014-0495-5

950 Zhang, D., Yang, Y., Li, M., Lu, Y., Liu, Y., Jiang, J., Liu, R., Liu, J., Huang, X., Li, G., & Qu, J.
951 (2022). Ecological Barrier Deterioration Driven by Human Activities Poses Fatal Threats to
952 Public Health due to Emerging Infectious Diseases. *Engineering*, *10*, 155–166.
953 https://doi.org/10.1016/j.eng.2020.11.002

954 Zhang, Z., Schwartz, S., Wagner, L., & Miller, W. (2004). A Greedy Algorithm for Aligning DNA
955 Sequences. *Https://Home.Liebertpub.Com/Cmb*, *7*(1–2), 203–214.
956 https://doi.org/10.1089/10665270050081478

957 Zinniel, D. K., Lambrecht, P., Harris, N. B., Feng, Z., Kuczmarski, D., Higley, P., Ishimaru, C. A.,
958 Arunakumari, A., Barletta, R. G., & Vidaver, A. K. (2002). Isolation and Characterization of
959 Endophytic Colonizing Bacteria from Agronomic Crops and Prairie Plants. *Applied and*
960 *Environmental Microbiology*, *68*(5), 2198–2208. https://doi.org/10.1128/AEM.68.5.2198-
961 2208.2002

962