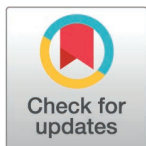


RESEARCH ARTICLE

# Diversification of an emerging bacterial plant pathogen; insights into the global spread of *Xanthomonas euvesicatoria* pv. *perforans*

Sujan Timilsina<sup>1</sup>✉, Fernanda Iruegas-Bocardo<sup>1</sup>✉, Mustafa O. Jibrin<sup>1,2,3</sup>✉, Anuj Sharma<sup>1,2b</sup>, Aastha Subedi<sup>1</sup>✉, Amandeep Kaur<sup>1</sup>, Gerald V. Minsavage<sup>1</sup>, Jose C. Huguet-Tapia<sup>1</sup>, Jeannie Klein-Gordon<sup>1,2c</sup>, Pragya Adhikari<sup>4</sup>, Tika B. Adhikari<sup>5</sup>, Gabriella Cirvilleri<sup>6</sup>, Laura Belen Tapia de la Barrera<sup>7</sup>, Eduardo Bernal<sup>8</sup>, Tom C. Creswell<sup>9</sup>, Tien Thi Kieu Doan<sup>10</sup>, Teresa A. Coutinho<sup>11</sup>, Daniel S. Egel<sup>9</sup>, Rubén Félix-Gastélum<sup>12</sup>, David M. Francis<sup>8</sup>, Misrak Kebede<sup>13</sup>, Melanie Lewis Ivey<sup>14</sup>, Frank J. Louws<sup>4,5</sup>, Laixin Luo<sup>15</sup>, Elizabeth T. Maynard<sup>16</sup>, Sally A. Miller<sup>14</sup>, Nga Thi Thu Nguyen<sup>10</sup>, Ebrahim Osdaghi<sup>17</sup>, Alice M. Quezado-Duval<sup>18</sup>, Rebecca Roach<sup>19</sup>, Francesca Rotondo<sup>14</sup>, Gail E. Ruhl<sup>9</sup>, Vou M. Shutt<sup>11,20</sup>, Petcharat Thummabenjapone<sup>21</sup>, Cheryl Trueman<sup>22</sup>, Pamela D. Roberts<sup>1,3</sup>, Jeffrey B. Jones<sup>1\*</sup>, Gary E. Vallad<sup>1,23\*</sup>, Erica M. Goss<sup>1,24\*</sup>



**1** Department of Plant Pathology, University of Florida, Gainesville, Florida, United States of America, **2** Department of Crop Protection, Ahmadu Bello University, Zaria, Nigeria, **3** Southwest Florida Research and Education Center, University of Florida, Immokalee, Florida, United States of America, **4** Department of Horticultural Science, North Carolina State University, Raleigh, North Carolina, United States of America, **5** Department of Entomology and Plant Pathology, North Carolina State University, Raleigh, North Carolina, United States of America, **6** Dipartimento di Agricoltura, Alimentazione e Ambiente, Sezione Patologia Vegetale, Catania, Italy, **7** Centro de Investigación en Alimentación y Desarrollo, Culiacán, Sinaloa, Mexico, **8** Department of Horticulture and Crop Science, The Ohio State University, Wooster, Ohio, United States of America, **9** Botany and Plant Pathology Department, Purdue University, West Lafayette, Indiana, United States of America, **10** Department of Plant Protection, College of Agriculture, Can Tho University, Can Tho, Vietnam, **11** Department Biochemistry, Genetics and Microbiology, Centre for Microbial Ecology and Genomics, Forestry and Agricultural Biotechnology Institute, University of Pretoria, Pretoria, South Africa, **12** Departamento de Ciencias Naturales y Exactas, Universidad Autónoma de Occidente, Unidad Regional Los Mochis, Los Mochis, Sinaloa, México, **13** Biotechnology Department, College of Biological and Chemical Engineering, Addis Ababa Science and Technology University, Addis Ababa, Ethiopia, **14** Department of Plant Pathology, The Ohio State University, Wooster, Ohio, United States of America, **15** Department of Plant Pathology, China Agricultural University, Beijing, China, **16** Department of Horticulture and Landscape Architecture, Purdue University, West Lafayette, Indiana, United States of America, **17** Department of Plant Protection, College of Agriculture, University of Tehran, Karaj, Iran, **18** Laboratorio de Fitopatología, Embrapa Hortaliças, Brasília-DF, Brazil, **19** Queensland Department of Agriculture and Fisheries, Brisbane, Queensland, Australia, **20** Department of Plant Agriculture, Ridgetown Campus, University of Guelph, Ridgetown, Ontario, Canada, **21** Department of Plant Science and Biotechnology, University of Jos, Jos, Nigeria, **22** Gulf Coast Research and Education Center, University of Florida, Wimauma, Florida, United States of America, **23** Division of Entomology and Plant Pathology, Faculty of Agriculture, Khon Kaen University, Khon Kaen, Thailand, **24** Emerging Pathogen Institute, University of Florida, Gainesville, Florida, United States of America

## OPEN ACCESS

**Citation:** Timilsina S, Iruegas-Bocardo F, Jibrin MO, Sharma A, Subedi A, Kaur A, et al. (2025) Diversification of an emerging bacterial plant pathogen; insights into the global spread of *Xanthomonas euvesicatoria* pv. *perforans*. PLoS Pathog 21(4): e1013036. <https://doi.org/10.1371/journal.ppat.1013036>

**Editor:** Eva H. Stukenbrock, CAU: Christian-Albrechts-Universität zu Kiel, GERMANY

**Received:** May 20, 2024

**Accepted:** March 12, 2025

**Published:** April 9, 2025

**Copyright:** © 2025 Timilsina et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data availability statement:** The data generated in this study, including raw read files and genome assemblies for strains sequenced, are deposited in NCBI under BioProject PRJNA941448. The sources of genome assemblies acquired from public databases are listed in [S1 Table](#).

✉ These authors contributed equally to this work.

✉a Current Address: Department of Entomology and Plant Pathology, Oklahoma State University, Stillwater, Oklahoma, United States of America

✉b Current Address: Horticultural Sciences Department and Gulf Coast Research and Education Center, University of Florida, Wimauma, Florida, United States of America

✉c Current Address: Agricultural Research Service, Crop Bioprotection Research Unit, Peoria, Illinois, United States of America

\* [emgoss@ufl.edu](mailto:emgoss@ufl.edu) (EMG); [gvallad@ufl.edu](mailto:gvallad@ufl.edu) (GEV); [jbjones@ufl.edu](mailto:jbjones@ufl.edu) (JBJ)

**Funding:** This research was supported in part by the United States Department of Agriculture's National Institute of Food and Agriculture awards 2015-51181-24312 (GEV), 2020-67013-31921 (EMG), and 2022-51181-38242 (GEV), by awards from the Ontario Agri-Food Innovation Alliance (UoG2015-2222), Ontario Tomato Research Institute, and the Fresh Vegetable Growers of Ontario to CT, and from the North Carolina Tomato Growers Association to TBA and FJL. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Abstract

Emerging and re-emerging plant diseases continue to present multifarious threats to global food security. Considerable recent efforts are therefore being channeled towards understanding the nature of pathogen emergence, their spread and evolution. *Xanthomonas euvesicatoria* pv. *perforans* (*Xep*), one of the causal agents of bacterial spot of tomato, rapidly emerged and displaced other bacterial spot xanthomonads in many tomato production regions around the world. In less than three decades, it has become a dominant xanthomonad pathogen in tomato production systems across the world and presents a compelling example for understanding diversification of recently emerged bacterial plant pathogens. Although *Xep* has been continuously monitored in Florida since its discovery, the global population structure and evolution at the genome-scale is yet to be fully explored. The objectives of this work were to determine genetic diversity globally to ascertain if different tomato production regions contain genetically distinct *Xep* populations, to examine genetic relatedness of strains collected in tomato seed production areas in East Asia and other production regions, and to evaluate variation in type III secretion effectors, which are critical pathogenicity and virulence factors, in relationship to population structure. We used genome data from 270 strains from 13 countries for phylogenetic analysis and characterization of type III effector gene diversity among strains. Our results showed notable genetic diversity in the pathogen. We found genetically similar strains in distant tomato production regions, including seed production regions, and diversification over the past 100 years, which is consistent with intercontinental dissemination of the pathogen in hybrid tomato production chains. Evolution of the *Xep* pangenome, including the acquisition and loss of type III secreted effectors, is apparent within and among phylogenetic lineages. The apparent long-distance movement of the pathogen, together with variants that may not yet be widely distributed, poses risks of emergence of new variants in tomato production.

## Author summary

Bacterial diseases pose significant threats to food security by reducing crop yield and increasing production costs. Managing these diseases is particularly challenging when pathogen populations are genetically diverse, rapidly evolving, and capable of long-distance dispersal. *Xanthomonas euvesicatoria* pv. *perforans* (*Xep*), one of the pathogens responsible for bacterial spot disease of tomato, exemplifies these challenges. Since its discovery in Florida in 1991, *Xep* has spread globally, affecting tomatoes in warm and wet production regions worldwide. Previous studies have indicated genetic diversity within *Xep*, suggesting the emergence of multiple lineages, but there was not a comprehensive global analysis. Here, we examined genome sequences of *Xep* strains from five continents and discovered extensive genetic diversity, including in genes important for virulence and breeding for resistance in tomatoes. In addition, we found that genetically similar strains were present on different continents, likely due to the international movement of contaminated seeds through global production chains. Our findings underscore the need for periodic monitoring of these pathogen populations and new approaches to effectively manage bacterial spot disease of tomato.

## Introduction/Main

Emerging and re-emerging plant diseases are a constant threat to global food security [1–3]. Bacterial plant pathogens cause some of the most intractable diseases of crops worldwide [4–7]. Novel emergence and re-emergence of bacterial diseases continue to be reported across the globe and are associated with an upsurge in efforts devoted to understanding the nature of pathogen emergence, spread, and evolution [8–17]. A bacterial plant pathogen that emerged in the last few decades and is of global epidemiological consequences is *Xanthomonas euvesicatoria* pv. *perforans* (*Xep*) [18] (syn. *X. perforans* [19, 20]), one of the causal agents of bacterial spot of tomato [21].

Bacterial spot disease of tomato affects all aboveground plant parts including leaves, stems, flowers and fruit. Under optimal environmental conditions, fruit lesions and/or extensive defoliation can dramatically limit marketable yields and pose a continuous challenge to tomato production [22–24]. Once epidemics are initiated, growers have limited management tools and have relied heavily on copper-based bactericides. However, reliance on copper compounds has led to widespread copper tolerance [25–33]. Alternative bactericides are often costly, provide insufficient control when the weather favors rapid disease development, and rarely improve yields. While historically four taxa have caused this disease, *Xep* has emerged rapidly and become a major player in tomato [21,30,34–41]. *Xep* was first reported in 1991 in Florida, USA [19] and is now found in all tomato production areas of the world, including regions with no history of the disease [42]. *Xep* has been isolated from tomato seed [19]; therefore, a likely hypothesis for new outbreaks of *Xep* is pathogen movement with seeds and planting materials [43].

Tomato production is characterized by a high seed replacement rate (99.3%), meaning that growers require seeds each season, which in turn requires large-scale seed production [44]. Tomato hybrid seed production is concentrated in geographic areas where environmental conditions minimize seed contamination by pathogens and seed production costs are low. These seed production regions supply hybrid seeds globally for commercial production of tomato fruits for the fresh market or for processing into tomato products (e.g., sauce, paste, and diced tomatoes). The long-distance movement of seeds poses a high risk for dissemination of seed borne pathogens to commercial tomato production areas. Seedlings are typically grown in transplant facilities and then transplanted into fields for the regional and international transplant markets, potentially amplifying and further disseminating seed-borne pathogens [45].

The success of *Xep* as a pathogen has been attributed to its production of bacteriocins against competing bacterial spot species, rapid genome evolution via recombination, and introduction of genes via horizontal gene transfer that contribute to fitness in tomato fields [46–52]. Distinct genetic lineages of *Xep*, each with unique patterns of allelic variation among core genes (genes present in all strains), were identified in fresh market and processing tomato production fields in the United States [47,50,51,53,54]. Additional lineages of *Xep* were found in Nigeria, Iran, Italy, and the Southwest Indian Ocean islands [42,48,55].

*Xep* strains, like other xanthomonads, acquire nutrients through colonization of susceptible hosts. The type III secretion system (T3SS) and type III effector (T3E) proteins are critical for suppression of host defenses and virulence by *Xep* [56]. Effector content varies among *Xanthomonas* species and distinct lineages of *Xep* have distinguishable effector content [26,48,50,57,58]. Strains of *Xep* isolated in the 1990s were unable to cause disease on pepper [59], but now strains of *Xep* are causing bacterial spot disease of pepper [50,58,60]. Host range expansion was attributed, in part, to loss of effectors that act as avirulence factors in pepper and other genomic changes as a result of recombination with other *Xanthomonas* lineages.

Effector variation can cause differences in disease epidemiology in addition to host range [49,57,61]. For example, wildtype strains with the acquired effector XopJ2 showed three times faster spread in the field than isogenic mutant strains without the effector [51].

Emerging pathogens can show limited genetic variation if they experienced a bottleneck during the ecological and evolutionary processes that often precede emergence (e.g., host jump or introduction event) [62]. *Xep* appears genetically diverse but it is not known how this variation is structured across global tomato production regions. The first objective of this work was to determine if different tomato production regions contain genetically distinct *Xep* populations. Second, we asked if there was evidence for long-distance pathogen dissemination, as would be indicated by genotypes shared among distant regions. Specifically, we obtained strains from tomato seed production areas in East Asia and asked if they resembled strains from fruit production fields elsewhere in the world, which would be expected if strains are being disseminated in seeds. Third, we estimated the timing of *Xep* population expansion relative to its first report in 1991. Finally, we evaluated T3E content and allelic variation in the context of geography and core genome variation. Overall, we found extensive genetic diversity within *Xep*; genetically similar strains in distant geographic regions, inclusive of seed production regions; evidence of diversification prior and subsequent to the first report of emergence; and lineage-specific T3E repertoires. Together, these results illustrate the capacity for this pathogen to rapidly evolve and strongly support the potential for intra- and intercontinental movement of pathogens in tomato production systems.

## Results

### *X. euvesicatoria* pv. *perforans* strains from seed and commercial fruit production areas

A total of 270 *Xep* genomes from 13 different countries – representing seed and fruit production – were used in this study (Table 1). We generated new genome sequence data for 153 strains of these strains and used published data for 117 (See accessions in S1 Table). *Xep* strains were differentiated from other tomato-pathogenic xanthomonads using a real-time qPCR assay that specifically amplifies the *hrcN* (*hrpB7*) gene in *Xep* [63] and inoculated on tomato cv. ‘Bonny Best’ to confirm pathogenicity. Strains from China, Thailand, and Vietnam were collected from seed production areas ( $n = 31$ ) and all other strains ( $n = 239$ ) were collected in commercial fruit production areas from Australia, Brazil, Canada, Ethiopia, Iran, Italy, Mexico, Nigeria, South Africa, and the United States. Within the US, strains were collected from seven different states in the Midwest and Southeast, including strains collected since 1991 from Florida.

### Genomic diversity in *X. euvesicatoria* pv. *perforans*

To examine genetic diversity in the core genome, we curated a set of 887 genes that were present in all 270 *Xep* genomes based on IMG/JGI gene annotation. The aligned sequence length of concatenated core genes was 617,855 bp, which contained 14,427 polymorphic sites after removing ambiguous nucleotides and any alignment gaps (S1 Data). Grouping strains by state within the United States and country elsewhere produced an FST [68] of 0.66. The values of Watterson’s  $\theta$  per site for the entire 617,854 bp alignment by geographic location, when represented by more than one strain, ranged from  $9.16 \times 10^{-6}$  (15 SNPs across the core gene alignment) to 0.00313 (5929 SNPs) (S2 Table). Nucleotide diversity (average number of differences per site, [69]) ranged from  $5.50 \times 10^{-6}$  to 0.00208. Both extremes in diversity came from the Midwestern U.S., Ohio and Indiana respectively (S2 Table). Tajima’s D [70] by geographic location ranged from  $-2.06$  to  $1.71$ , but many locations had low sample sizes (S2 Table).

**Table 1. *Xanthomonas euvesicatoria* pv. *perforans* strains used in this study.**

Country	Locality	Year	Strain (original name, if applicable)
Australia [57]	Queensland	2015	Aus3, Aus7, Aus14
		2016	Aus5, Aus10, Aus11
		2017	Aus1, Aus15, Aus16
Brazil [39]	São Paulo	2011	Bzl1 (2011-107), Bzl2 (2011-132)
	Goiás	2012	Bzl3 (2012-08)
	Goiás, São Paulo	2013	Bzl5 (2013-16), Bzl6 (2013-42)
	Goiás, Minas Gerais	2014	Bzl7 (2014-10), Bzl8 (2014-17)
	Minas Gerais	2015	Bzl10 (2015-53), Bzl11 (2015-56)
	Goiás	2016	Bzl13 (2016-08)
	Goiás	2017	Bzl14 (2017-21)
Canada	Ontario	2016	4A, 4D, 12A, 14A
China		2016	CHI-3, CHI-5, CHI-6, CHI-7, CHI-8, CHI-10, CHI-12, CHI-15, CHI-18
Ethiopia [36]		2011	ETH5, ETH11, ETH21, ETH25, ETH33
Iran [41]		2013	K41, F210, F215, TOM801, TOM816
Italy [64]		2011	1P6S1, 2P4S1, 2P4S1D, 2P6S1, 1P4S1D
Mexico			Mexico-1, Mexico-3, Mexico-LT1, Mexico-LT3, Mexico-LT5
Nigeria [37,65]		2014	NI-1, NI-2, NI-4, NI-7, NI-12, NI-13
		2015	KS3, KS5, KS9, KS28
South Africa	Pretoria		X2-B14, X10-B85, X59-BD1351, X47-BD167
Vietnam			SEA-3, SEA-5, SEA-21, SEA-23
Thailand		2016	THA-8, THA-14, THA-40, THA-45, THA-54, THA-72, THA-81A, THA-100, THA-112, THA-116, THA-119, THA-120, THA-126, THA-127, THA-128, THA-132, THA-135, THA-157A
United States	Alabama [66]	1996	Xp1861
	Indiana [28]	2016	16-1165A1, 16-1181-2, 16-1182A, 16-1184A, 16-1187A, 16-1205A, 16-1402A, 16-974C, 16-990A, 16-990C
		2014	14-463-1A
	Florida [19,20,43,47,58,67]	1991	XV0938, Xp91-118, Xp894, Xp909, Xp1183
		1992	Xp1118, Xp1144
		1993	Xp1241, Xp1268, Xp1275
		1994	Xp1550, Xp1564
		1995	Xp1797, Xp1805
		1996	Xp1856
		1997	Xp1912
		1998	Scott-1, Xp1920
		2006	Xp1-5, Xp1-6, Xp3-12, Xp3-15, Xp3-16, Xp3-8, Xp4-20, Xp5-14, Xp5-6, Xp5-9, Xp7-12, Xp8-16, Xp9-5, Xp10-13, Xp11-2, Xp15-11, Xp17-12, Xp18-15
		2007	Xp4B
		2010	Xp2010
		2011	GEV485
		2012	GEV839, GEV872, GEV893, GEV904, GEV909, GEV915, GEV917, GEV936, GEV940, GEV968, GEV993, GEV1001, GEV1026, GEV1044, GEV1054, GEV1063
		2013	TB6, TB9, TB15

(Continued)

Table 1. (Continued)

Country	Locality	Year	Strain (original name, if applicable)
		2015	GEV2047, GEV2048, GEV2049, GEV2050, GEV2052, GEV2055, GEV2058, GEV2059, GEV2060, GEV2063, GEV1989, GEV1991, GEV1992, GEV1993, GEV2004, GEV2009, GEV2010, GEV2011, GEV2013, GEV2015, GEV1911, GEV1912, GEV1913, GEV1914, GEV1915, GEV1916, GEV1917, GEV1918, GEV1919, GEV1920, GEV1921
		2016	GEV2065, GEV2067, GEV2072, GEV2087, GEV2088, GEV2089, GEV2097, GEV2098, GEV2099, GEV2108, GEV2109, GEV2110, GEV2111, GEV2112, GEV2113, GEV2114, GEV2115, GEV2116, GEV2117, GEV2118, GEV2119, GEV2120, GEV2121, GEV2122, GEV2123, GEV2124, GEV2125, GEV2126, GEV2127, GEV2128, GEV2129, GEV2130, GEV2132, GEV2133, GEV2134, GEV2135
	Louisiana [27]	2013	mli-2
	North Carolina [30]	2015	NC-14, NC-47, NC-67, NC-101, NC-112, NC-204
		2016	NC-242, NC-252, NC-282, NC-289, NC-350, NC-373, MRS-30P-011
	South Carolina [43]	2016	GEV2407, GEV2408, GEV2384, GEV2388, GEV2389, GEV2390, GEV2391, GEV2392, GEV2393, GEV2396, GEV2397, GEV2399, GEV2400, GEV2403, GEV2410, GEV2420
	Ohio [53]	2017	SM-1806, SM-1807, SM-1808, SM-1809, SM-1810, SM-1811, SM-1812, SM-1813, SM-1814, SM-1815, SM-1828, SM-1829, SM-1830, SM-1831

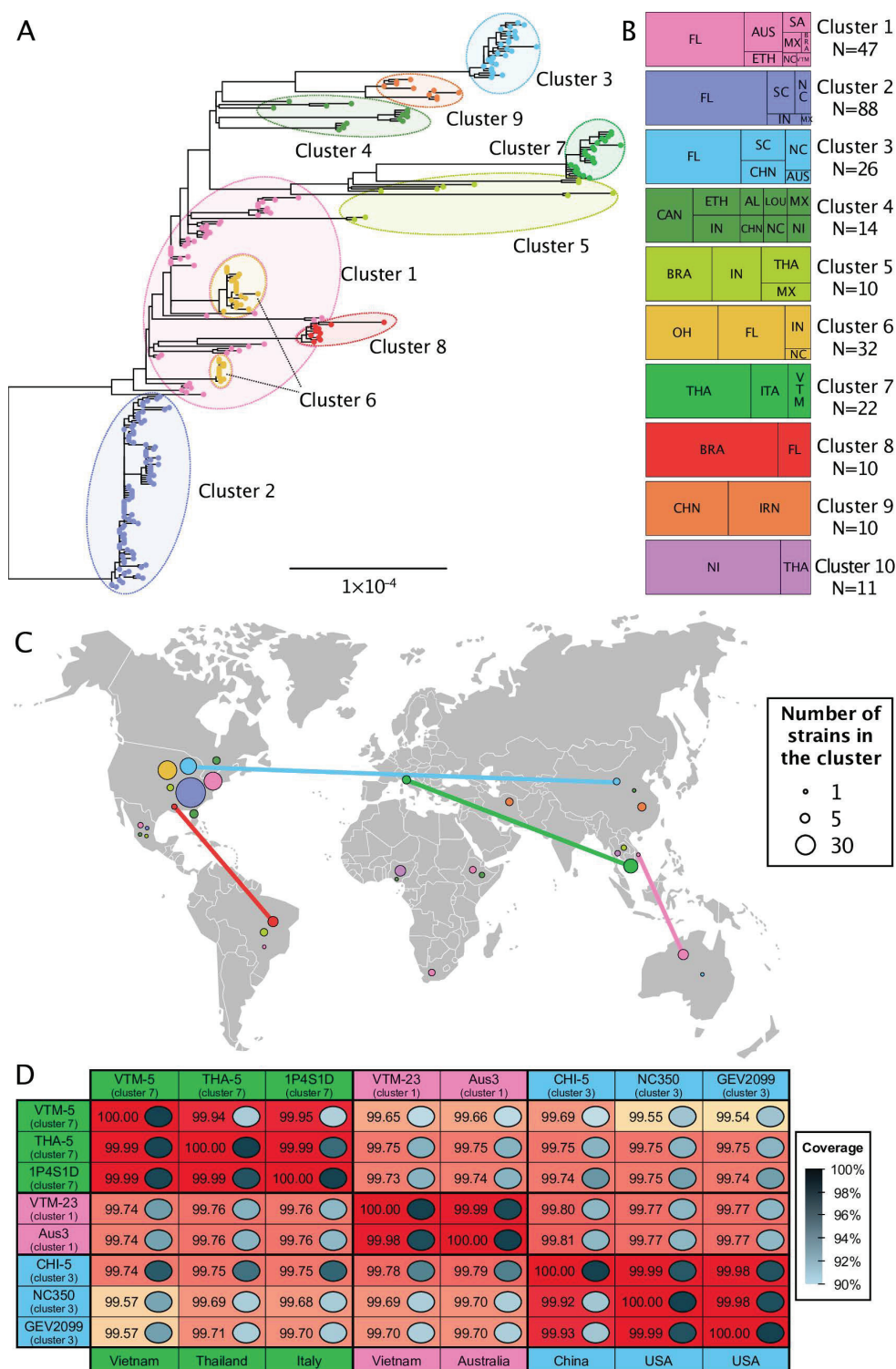
<https://doi.org/10.1371/journal.ppat.1013036.t001>

Maximum likelihood phylogenetic analysis of core SNPs revealed diversifying lineages of *Xep* (Fig 1) and an especially diverged lineage of 11 strains from Nigeria and Thailand (S1 Fig) that included a previously defined atypical strain – NI1 – from Nigeria [48]. After excluding the 11 atypical strains due to their phylogenetic divergence, ClonalFrameML [71] estimated an overall ratio of recombination rate to mutation rate ( $R/\theta$ ) of 0.60, with recombination causing approximately seven times more base changes than mutation ( $\delta = 231$ ;  $\nu = 0.05$ ). There were an estimated 221 recombination events that affected more than 96 Kbp in terminal branches and 494 recombination events detected in internal branches encompassing 190 Kbp.

To summarize population structure based on core gene SNPs, we used hierBAPS [72], which assigned individual strains to 9 clusters using allele frequencies (Fig 1 and S1 Table). This analysis did not include the 11 highly diverged strains from Nigeria and Thailand, which we designated as cluster 10.  $F_{ST}$  among clusters was 0.80. In some cases, clusters corresponded to phylogenetic lineages, including clusters 2, 3, 7, 8, and 9 (Fig 1). The remaining clusters were polyphyletic, encompassing multiple diverged clades or individual strains. Nucleotide diversity within clusters ranged from  $2.23 \times 10^{-5}$  for cluster 8 to 0.0011 for cluster 5, while cluster 3 had the highest diversity for a monophyletic cluster ( $6.98 \times 10^{-4}$ ) (S2 Table). Watterson's  $\theta$  per site was lowest for cluster 8 ( $3.57 \times 10^{-5}$ ) and highest for cluster 6 (0.0018), cluster 3 (0.0013), and cluster 5 (0.0011) (S2 Table). Analysis of presence-absence gene variation in the pangenome showed that polyphyletic clusters 1, 4, and 5 had the most variation in gene content (S2 Fig).

### Geographic distribution of *X. euvesicatoria* pv. *perforans* core gene clusters

Cluster 1 encompasses genetically diverse strains from seven countries, including most of the strains from Australia, all four strains from South Africa, and one strain from Southeast Asia (Fig 1B). All USA strains assigned to cluster 1 were isolated in or before 2006 from Florida except for one strain from North Carolina. Cluster 2 contains 88 strains from the United States and one from Mexico, while Cluster 3 includes strains isolated from Florida, North



**Fig 1. Population structure of *Xanthomonas euvesicatoria* pv. *perforans* strains collected from tomato production regions.** (A) Maximum likelihood phylogenetic tree of 259 *X. euvesicatoria* pv. *perforans* strains constructed with nucleotide sequences from 887 core genes, corrected for recombination by ClonalFrameML. Tips are colored according to clusters identified by hierBAPS. Nucleotide alignment is available as [S1 Data](#). (B) Distribution of 270 strains in each cluster by country or state of collection. Strains designated as cluster 10 (n=11) were genetically distant and

excluded from the tree and hierBAPS analysis (see [S1 Fig](#)). N indicates total number of strains in each cluster. Geographic abbreviations are as follows: AUS – Australia; BRA – Brazil; CAN – Canada; CHN – China; ETH – Ethiopia; FL – Florida, USA; IN – Indiana, USA; IRN – Iran; ITA – Italy; LOU – Louisiana, USA; MX – Mexico; NC – North Carolina, USA; NI – Nigeria; AL – Alabama, USA; OH – Ohio, USA; SA – South Africa; SC – South Carolina, USA; THA – Thailand; VTM – Vietnam. (C) Map showing distribution of clusters by country of collection. Lines show instances of strains with high core gene sequence identity that were collected in different countries ([S3 Table](#)). Base layer of the map is courtesy of Eurostat (<https://ec.europa.eu/eurostat/web/gisco>). (D) Pairwise comparison of whole genome average nucleotide identity (ANI) confirmed high identity between strains isolated from different continents. For each comparison, genome coverage is shown by grayscale in boxes, scale shown to the right. Values for each comparison are for genomes in rows when compared to genomes in columns. See [S3 Table](#) for additional ANI output.

<https://doi.org/10.1371/journal.ppat.1013036.g001>

and South Carolina, China, and Australia. Cluster 4 encompasses multiple lineages of strains from the United States, Canada, Ethiopia, China, and Nigeria. Cluster 5 is polyphyletic with diverged strains from the United States, Mexico, Brazil, and Thailand. Cluster 6 was isolated only within the United States from Florida, Indiana, North Carolina, and Ohio. Cluster 7 is a monophyletic group of strains from Southeast Asia and Italy. Cluster 8 is another monophyletic group found only in Brazil and Florida. Cluster 9 includes two clades of strains, one from China and the other from Iran and Nigeria. Cluster 10 comprises the atypical strains from Nigeria and similar strains from Thailand. Most countries contained strains from more than one core gene cluster ([Fig 1C](#)).

Clusters 1, 3, 4, 5, 7, 9, and 10 contain strains isolated from both seed production and commercial fruit production regions, whereas strains in clusters 2, 6, and 8 were only isolated from commercial fruit production regions. Some strains found on different continents were nearly identical in core gene sequences with very high average nucleotide identity ([Fig 1D](#)). Strains in cluster 1 from Australia differed by 6 to 10 SNPs in more than 617 Kbp of core gene sequence from strain VTM-23 from Vietnam. Whole genome pairwise average nucleotide identity (ANIb) between VTM-23 and Aus3 was 99.99% compared to ANIb values ranging from 99.65 to 99.81 for comparisons to genomes representing other clusters ([Fig 1D](#)). Strains from the USA had up to 99.87 ANIb with strains from Australia and Vietnam ([S3 Table](#)). A different strain from Vietnam, VTM-5 in cluster 7, had as few as four SNPs in the core genome when compared to strains from Italy and ANIb of 99.95% to Italian strain 1P4S1D ([Fig 1D](#)). Likewise, strains collected in a seed production region in China had ANIb up to 99.99% with strains from Florida and North Carolina. We also found similar strains between Brazil and USA, for example Bzl-10 (Minas Gerais) and Xp3-8 (Florida) had greater than 99.9% ANI ([S3 Table](#)). Other strains were similar between countries in core genes only after correction for recombination.

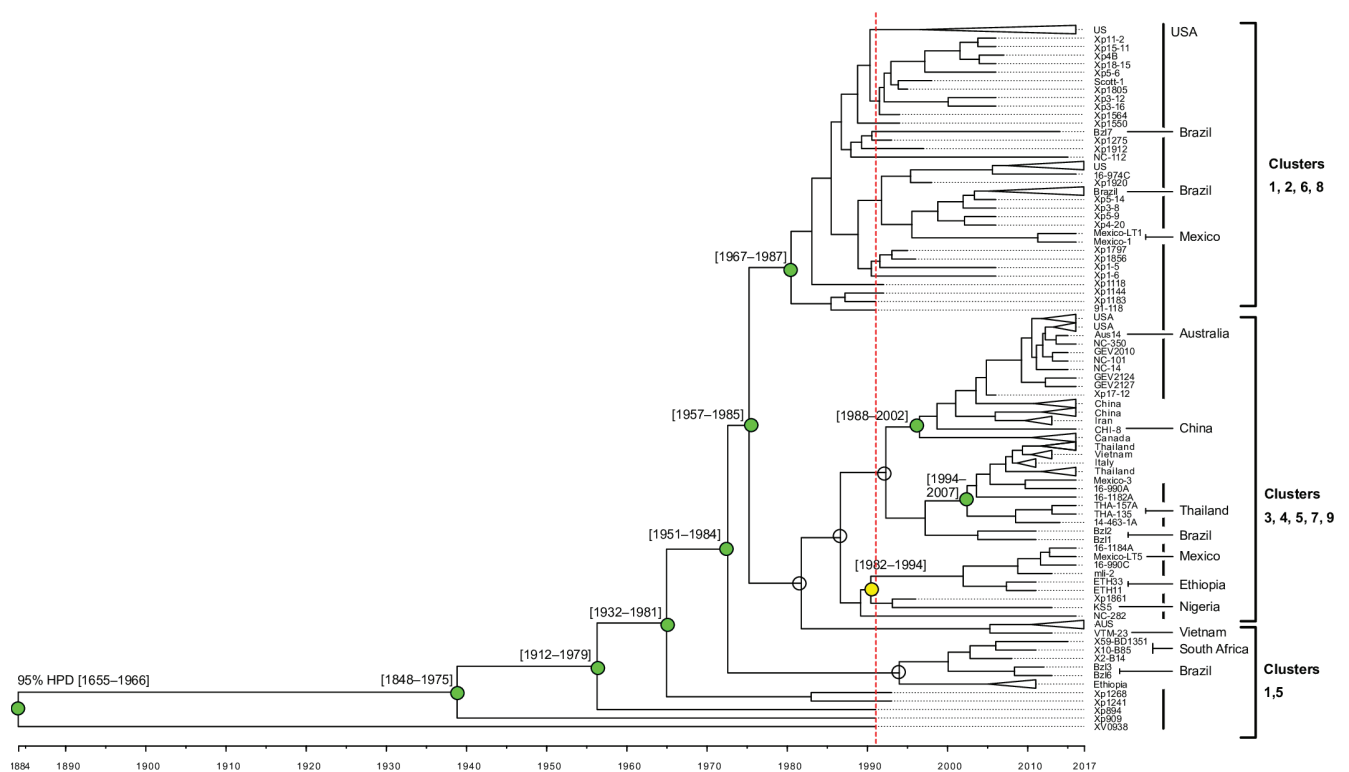
### Timing of *X. euvesicatoria* pv. *perforans* lineage emergence

We used the years of strain collection to estimate the timing of diversification of our sample of *Xep*, excluding the cluster 10 strains. We inferred dated phylogenies using whole genome alignments with inferred recombinant sites removed by Gubbins [73]. Due to recombination with other *X. euvesicatoria* lineages, we did not include an outgroup ([S1B Fig](#); [48]). Sampling year was significantly correlated with root-to-tip distance ( $R^2 = 0.20$  for the whole genome alignment,  $P < 1 \times 10^{-4}$ , [S3 Fig](#)). The root inferred by the BactDating R package [74] was placed between strains isolated in Florida in 1991. The most recent common ancestor (MRCA) of all strains was dated to 1884 (95% HPD: 1655–1966). Notably, strains that were isolated in the early 1990s, when *Xep* was first detected in U.S. tomato production [19,37], represented multiple lineages ([Fig 2](#)). The MRCA of the clade representing core gene clusters 1, 2, 6, and 8 (including strains from USA, Brazil, and Mexico) was dated to 1980 (95% HPD: 1967–1987). A major clade, encompassing strains in clusters 3, 4, 5, 7, 9, which were collected in Africa,

the Americas, Asia, Australia, and Europe, did not have a significant temporal signal across the clade. We repeated the analysis with BEAST, which inferred a different rooting. The tree inferred by BEAST placed the root between two strains isolated in 2011 from Brazil and all other strains (S4B Fig). The MCRA of the BEAST tree was dated to 1868 (95% HPD: 1862–1919), which was similar to the root date estimated using BactDating (S4 Fig).

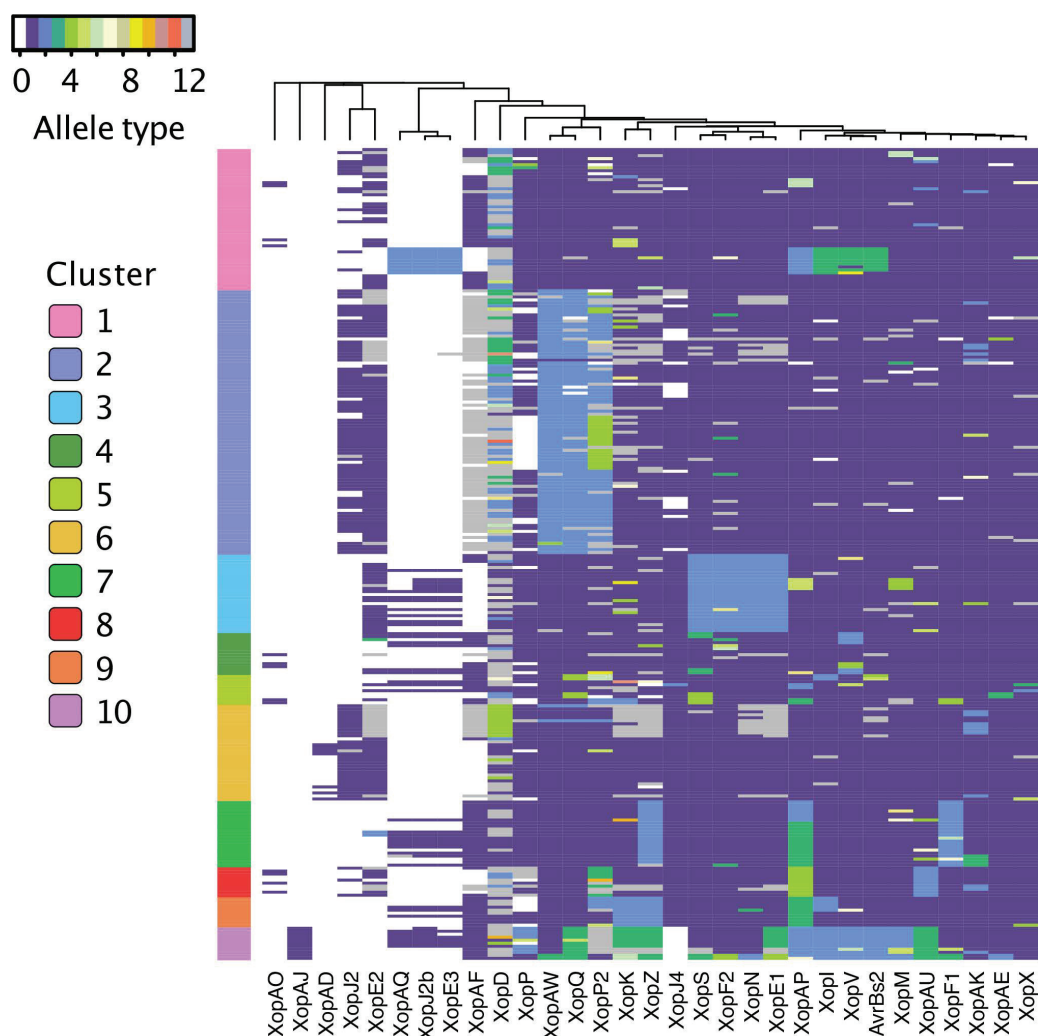
### Type III effector content

We detected 32 predicted type III effector genes in our collection of 270 strains (S5 Fig and S4 Table). The diversity in amino acid sequences of each predicted effector ranged widely from a single conserved allele to 8 or more alleles per gene (Fig 3). None of the effectors were present and intact in 100% of our genomes, partly due to our analysis of draft genomes. The following fourteen effector genes were present in more than 95% of strains and can be considered “core effectors”: *avrBs2*, *xopF1*, *xopF2*, *xopI*, *xopM*, *xopQ*, *xopS*, *xopV*, *xopX*, *xopAE*, *xopAK*, *xopAP*, *xopAU*, and *xopAW*. The genes for *xopD*, *xopE1*, and *xopN* were present in some form in all genomes but more than 5% of strains contained a contig break within the gene. A closer examination of *xopD* by PCR and Sanger sequencing showed this to be an assembly issue due to the repeats within the gene. Effectors at low frequency in our *Xep* strains (<25%) were



**Fig 2. Dated phylogeny of 259 *X. euvesicatoria* pv. *perforans* strains.** BactDating analysis estimated an approximately 130-year history for *Xep* strains in core gene clusters 1 through 9 (Fig 1). Red dotted line indicates the first documented isolations in 1991. Internal nodes were collapsed for clades containing strains from a single country with branch tips indicating country or strain (for full tree see S4 Fig). Bold vertical lines to the right of tip labels indicate strains from USA; other countries are labeled. Temporal signal was assessed using PhyloStems and results are shown for major nodes (for full results see S3 Fig). Empty circles indicate no significant temporal signal. Colored circles indicate nodes with statistically significant temporal signal based on adjusted  $R^2$  values: green – 0.13–0.19; yellow – 0.45. The 95% highest posterior density (95% HPD) of date estimates for major nodes with significant temporal signals are shown in brackets.

<https://doi.org/10.1371/journal.ppat.1013036.g002>



**Fig 3. Variation in type III effectors (Xop proteins) in *Xanthomonas perforans*.** Type III effectors are in columns and 270 *X. euvesicatoria* pv. *perforans* strain in rows. Effector status is shown by allele type: absence is indicated by allele type 0 (white), while the most frequent allele observed when the effector is present is allele type 1 (purple), second most frequent is allele type 2 (blue), and so on. Putative pseudogenized effectors are shown as allele 13 (gray). The order of columns was determined by hierarchical clustering analysis, placing similarly distributed effectors adjacent to each other. Genomes showing BLAST hits to TAL effector(s) are indicated in [S3 Table](#) and not shown in heatmap. *X. euvesicatoria* pv. *perforans* strains (rows) are organized by core gene cluster.

<https://doi.org/10.1371/journal.ppat.1013036.g003>

*xopE3*, *xopAD*, *xopAJ*, *xopAO*, and *xopAQ*. Transcription activator-like (TAL) effectors typically do not assemble in draft genomes due to their characteristic repeat sequences, but there were BLAST hits to previously described TAL effectors in 65 strains. We Sanger sequenced the TAL effector gene in strain 2P6S1 collected in Italy (NCBI accession number OQ588696), which confirmed that it had the same repeat variable diresidues as PthXp1 reported in *Xep* strains from Alabama [50]. The strain isolated in Louisiana, USA was previously reported to have AvrHah1 [16].

The T3E effector XopAF (AvrXv3), which is targeted by the tomato resistance gene *Xv3* [75], was missing or pseudogenized in 64% of strains. Most strains examined from the United States did not have a complete copy of this gene, whereas it was intact in many

strains collected in Asia and Africa. The gene for XopJ4 (AvrXv4), recognized by resistance gene *RXopJ4* from *S. pennellii* [76], was present in 88% of strains and absent in all cluster 10 strains and 19 of 88 cluster 2 strains. XopJ2 (AvrBsT), which elicits an HR in pepper but increases virulence in tomato [49], was present in less than half of strains examined (43%) and overwhelmingly in strains from the United States. A homolog of XopJ2, recently designated XopJ2b [77], was present in 50 strains, including two strains from Australia that carried both copies of XopJ2 (S4 Table).

We tested for evidence of positive selection in T3E by estimating synonymous and non-synonymous (dN/dS) substitution rates using a Bayesian approach for detecting pervasive selection (FUBAR, [78]) and maximum likelihood approach for detecting episodic selection (MEME, [79]). We found evidence of pervasive positive selection affecting at least one amino acid in AvrBs2, XopD, XopE1, XopF2, XopK, XopM, XopP and its paralog XopP2, XopQ, XopS, and XopAQ (S4 Table). We found evidence of episodic selection affecting at least one amino acid in XopF2, XopK, XopP, XopP2, XopQ, XopV, and XopAP (S4 Table).

We defined the effector profile of each strain as the predicted presence or absence of each effector and its allelic state, excluding TAL effector hits. Grouping effector profiles according to core gene cluster revealed that allelic variation of effectors often paralleled core genome variation (Fig 3). For example, particular alleles of effectors XopAW, XopQ, and XopP2 were mostly limited to strains in cluster 2. Cluster 3 strains carried unique alleles for effectors XopF2, XopS, XopN, and XopE1, and strains from cluster 7 shared unique alleles for effectors XopF1 and XopZ. Strains from highly diverged cluster 10 had rare alleles in many effectors, and it was the only cluster in which effector XopAJ was found (Fig 3). To visualize variation among strain effector profiles independent of core gene clusters, we transformed dissimilarities between profiles into distances represented in a two-dimensional plot and defined eight effector profile clusters (S6A Fig). Cluster A was characterized by a lack of low frequency effectors and contained 188 strains from 11 of 13 countries (S6 Fig). The remaining effector clusters were defined by the presence of one to three low frequency effectors (S7 Fig). While most effectors were found in multiple countries and continents (S6 Fig), populations in Brazil, Ethiopia, Nigeria, Thailand, South Africa, and the United States contained low frequency effectors that were not widely distributed.

### Copper resistance genes

Xanthomonads, including *Xep*, have acquired genes that confer copper tolerance, likely in response to exposure to copper-based bactericides [33,80–82]. In *Xep*, copper tolerance is conferred by an operon containing the copper resistance genes *copA* and *copB*, and regulator *copL* (*copLAB*) [81]. BLAST analysis showed that these genes were present in 73% of the genomes in our sample (S5 Table). Copper resistance genes are prevalent in the USA; only the genomes from strains isolated from Florida in the early 1990s and a strain from Louisiana lacked *copLAB*. The genes were also missing in the genomes of a few strains from Australia (1), Brazil (2), Ethiopia (2), Mexico (1), and Vietnam (2). In contrast, the genes were absent in all genomes of all strains from Nigeria, China, Iran, Italy, and Thailand.

### Discussion

Emerging plant pathogens have the potential for global outbreaks, exacerbated by complex trade networks. Hybrid tomato production relies on international breeding and production chains with a global network to deliver seeds to growers. Global trade associated with vegetable seed production provides a pathway for global spread of pathogens, with quantities traded that challenge even strong phytosanitary measures [83, 84]. Over 100 countries import seeds of tomatoes and other vegetables; for example, 11.7 million kg of vegetable seed were

imported to the USA in 2019, with China being the biggest supplier at 2.4 million kg [85]. *Xanthomonas* species can infest pepper and tomato seed [86], and *Xep* has been isolated from tomato seed [19,87], supporting the hypothesis that seeds can be a source of inoculum for bacterial spot outbreaks [88, 89]. Thirty years after its first report, *Xep* has been identified in tomato production areas around the world [21]. Our results showed extensive genetic diversity in the pathogen, but also genetically similar strains in distant tomato production regions. Furthermore, we found genetically similar strains in seed production and fruit production regions on different continents, as would be expected if the pathogen was being moved in shared production chains. Dated phylogenies indicate multiple waves of diversification of the *Xep* population, before and since its first detection in 1991. Variation in gene content confirms that *Xep* acquired and lost type III effectors during its diversification, which will continue to challenge sustainable management of tomato bacterial leaf spot [49,66].

Using our broad strain collection, we found *Xep* variants in seed production regions in Asia that were previously reported in Australia, Italy, Nigeria, and the United States [43,47,48,57]. Strains from Italy were nearly identical in core genes and very similar in accessory genomes to strains collected from Thailand and Vietnam (cluster 7), both major seed production regions. The atypical bacterial spot strains from Nigeria, recently designated as race T5 [24], were genetically similar to strains from Thailand (cluster 10). Cluster 10 strains may encompass one or more new subspecific taxa within *X. euvesicatoria*. A recently described variant of *Xep* in Florida [cluster 3; [43,47]], which was also found in Australia [57], was similar in the core genome to strains found in China (cluster 3); however, these strains showed divergence in the pangenome, consistent with accessory genome evolution in emergent populations. Beyond previously described variants, we found strains in Iran that were closely related to strains from China (cluster 9); multiple instances of genetic similarity between strains from North America and Ethiopia (clusters 1 and 4); and highly similar strains shared between USA and Brazil (cluster 8), USA and Australia (clusters 1 and 3), and between Australia and Vietnam (cluster 1). Given the variation of *Xep* across our sample, genetic similarity in core genes and gene content across continents is strong evidence of international dissemination. Genetically similar strains of bacterial spot pathogens *X. euvesicatoria* pv. *euvesicatoria* and *X. hortorum* pv. *gardneri* collected from different continents similarly suggest intercontinental dissemination in tomato and pepper seed [15,58,90]. Whole genome analysis of *X. hortorum* pv. *pelargonii* strains from a 2022 epidemic of bacterial blight of geranium in the USA showed zero to seven chromosomal SNPs among isolates of the emergent strain that was distributed to multiple states in plant cuttings [9,91].

Other *Xep* genotypes indicated a more limited distribution. We did not find core gene cluster 2 strains in the seed production regions sampled (China, Thailand, Vietnam), while this lineage was highly represented in our USA sample. To date, strains in this cluster have been found only in the southeastern and midwestern USA [43,47,50,53,54] and Mexico. Seedling nurseries in the southeast USA produce tomato transplants for growers in multiple states. Interstate movement of strains on seedlings is likely responsible, at least in part, for disseminating genetically similar strains to different states [43,53]. We previously reported extensive recombination with *X. euvesicatoria* pv. *euvesicatoria* in cluster 2 strains [47] and this cluster had a diverse accessory genome. Additional research is needed to determine the population dynamics and genetic mechanisms that underly the diversity in this cluster.

The *Xep* strains we examined from the USA were assigned to core gene clusters 1, 2, 3, 4, 5, 6, and 8, representing several distinct genetic lineages. The US sample had an overall negative Tajima's D value. A negative value of Tajima's D across the USA sample suggests an abundance of low frequency haplotypes, which is consistent with low frequency clonal lineages or accumulation of mutations within lineages. Tajima's D varied from positive to negative within

individual states and in samples from other countries, although values for small sample sizes should be interpreted with caution.

To better understand the initial emergence of *Xep*, we used calibrated phylogenies to examine the timing of lineage divergence. International trade in F1 hybrid tomato seed surged in the second half of the 20<sup>th</sup> century, after the first hybrid tomato cultivars were released by 1940 [92, 93]. There was a 300-fold increase in hybrid tomato seeds exported from Asia between 1962 and 1977 [94, 95] and subsequent rapid growth in tomato production. Our analyses estimate the most recent common ancestor of our sample to ~150 years ago, while the major ancestral lineages diverged during or after the early expansion in the hybrid seed trade. We hypothesize that the emergence and geographic distribution of lineages may be associated with the multinational structure of tomato breeding and seed production, in which parental lines and geographic locations of seed production change over time [96].

Bacterial spot is a destructive disease in areas where tomatoes are grown under humid conditions and growers in the USA have relied heavily on copper bactericides to manage this disease. In response, *Xep* strains have developed copper tolerance [26]. Most strains isolated from Florida in the 1990s lacked the *copLAB* genes, but they are now common in strains collected in the USA [97]. Recent studies of Florida strains found that these copper resistance genes are more frequently present on the chromosome than on a plasmid, suggesting selection for vertical inheritance of copper tolerance [33,97]. In contrast, strains from other countries lacked copper resistance genes, indicating little or no local selection for the acquisition of *cop* genes.

Type III effectors are important members of *Xanthomonas* genomes given their roles in pathogenicity and virulence. We found up to 16 putative core effector genes, most of which exhibited allelic variation. The impact of allelic variation in *Xep* effectors on pathogen fitness, if any, is unknown. Signatures of positive selection on some genes indicate past or ongoing fitness impacts. Low frequency effectors were found across core gene clusters, suggesting acquisition of new effectors and their exchange among *Xep* lineages. For example, some strains in clusters 3, 4, and 5 from the United States, Canada, and Mexico carried the same alleles of low frequency effectors XopAQ and XopE3 as strains from Asia, Nigeria, and Italy. BLAST analysis suggested the geographically widespread presence of transcription activation-like (TAL) effectors in *Xep*. Both TAL effectors described in *Xep*, *avrHah1* and *pthXp1*, are associated with increased disease severity on tomato [16,50]. Acquisition of T3Es could increase the fitness of *Xep* relative to other bacterial spot pathogens and cause more damaging disease outbreaks [49,51,66].

The release of new plant varieties that carry disease resistance genes can have dramatic effects on pathogen population structure due to selection to overcome host resistance [98–100], and we have previously reported on the loss of function of effector AvrXv3 (XopAF) across lineages [30,53,66]. Examination of T3E content at a global scale puts variation previously observed in Florida into a larger context. XopAF was present in strains collected in the 1990s (cluster 1), but absent or non-functional in most strains from Florida, Indiana Ohio, and North Carolina, USA [26,28,30,53,66]. Here, we found that *xopAF* was intact in many strains from seed production areas, but confirmed that it is no longer a viable target for resistance in the USA. The effector XopJ4 is a potential resistance target [66] based on its recognition by *RXopJ4* from *S. pennellii* [76], but Klein-Gordon et al. [26] reported that it was missing from 3.2% of Florida strains collected in 2017 and, here, we found that it was absent in one North Carolina and 20 Florida, USA strains. All strains collected outside the USA contained *xopJ4*, except for cluster 10 strains. Another XopJ family member, *xopJ2*, is a virulence factor in tomato [49,51]. This T3E is common in North America, particularly in cluster 2 and 6 strains, but absent or infrequently detected in *Xep* populations elsewhere. An alternative form

of this effector, recently described as XopJ2b [77], is more common in strains from outside North America. Resistance to strains with XopJ2 (XopJ2a) is conferred by Ptr-1 from the wild relative of tomato *Solanum lycopersicoides* [101], which may also be effective against XopJ2b [77]. Alternative resistance strategies include transgenic *Bs2* [59] and CRISPR editing of the *bs5* homologs in tomato [102], which are expected to provide resistance to many or all *X. perforans* strains, respectively.

In summary, we found strong evidence for intercontinental movement of *Xep*, consistent with the international nature of tomato breeding and hybrid tomato seed production. We also found notable diversity in our global sample of *Xep*, including in seed production regions, and multiple variants of *Xep* that do not appear to be widely distributed. Our results also suggest that continued monitoring of bacterial spot pathogens is warranted to identify emerging lineages that may respond differently to disease management with copper-based products and effector-targeted host resistance. These findings also raise questions regarding the degree of genetic variation seen in *Xep*, such as the evolutionary genetic mechanisms that are responsible and the effects of different genetic variants on epidemiology. The genomic diversity of *Xep* in seed and fruit production regions creates the opportunity for recombination among strains and dissemination of high fitness variants of *Xep*.

## Materials and methods

### Bacterial strains, genome sequencing, and assembly

*Xep* strains were collected from 13 different countries (Tables 1 and S1). Strains from the United States were collected from seven states between 1991 to 2016 and comprised 181 strains. The remaining 89 strains were collected from Canada, Mexico, Brazil, Italy, Ethiopia, Nigeria, South Africa, Iran, China, Thailand, Australia, and Vietnam. Strains from China, Thailand, and Vietnam were collected from fields designated for production of tomato seed for the global market. Strains from Brazil were obtained from both staked fresh-market and processing tomato commercial fields. Strains from Italy were isolated from tomato pith in greenhouse tomato showing wilting symptoms [64,103]. Strains from South Africa were collected from commercial seed lots. Strains from Nigeria were obtained from fields cultivated for both subsistence and commercial purposes. Strains from other countries were collected from fields designated for commercial fruit production.

A total of 270 *Xep* genome sequences were used during this study (S1 Table). Draft and whole genomes of 117 strains were generated and published previously [43,47,48,57,58,67,103]. The remaining 153 strains were sequenced for this study using Illumina platforms. Genomic DNA was extracted from single colony cultures grown for 24-hr in nutrient broth using the Wizard Genomic DNA Purification Kit (Promega, Chicago, IL) following manufacturer instructions. Genomic libraries for sequencing were prepared using the Nextera DNA library preparation kit from Illumina (Illumina, San Diego, CA). Sequencing was performed at the Interdisciplinary Center for Biotechnology Research, University of Florida, using an Illumina MiSeq to generate 250 bp paired end reads for each strain. Additional genomic sequence data were generated for five strains for the ANI analysis (S1B Table). Genomic DNA was extracted using the above methods except that extracted genomic DNA was sent to SeqCenter (Pittsburg, PA) for sequencing with Illumina NovaSeq 6000, producing 150 bp paired end reads.

Raw reads were trimmed of adapters and paired with Trim Galore (<https://github.com/FelixKrueger/TrimGalore>) [104], then assembled into contigs with Spades version 3.10.1 [105], with k-mers 21, 33, 55, 77, 99, and 127 with read error correction and "--careful" switch. Reads were then aligned to the assembled contigs using Bowtie 2 v. 2.3.3 [106]. Inconsistencies

were identified and polished using Pilon [107]. Contigs smaller than 500 bp and with less than 2.0 k-mer coverage were filtered out. Quality of genomes was assessed with CheckM [108]. Assembled genomes were annotated using the IMG/JGI platform [109]. The genome data generated for this study are available in NCBI BioProject PRJNA941448.

### Core gene phylogeny

In a previous study, we defined a set of 1,356 ‘core genes’ from 58 genomes of *Xep* strains isolated from Florida [47]. The core genes were determined based on amino acid sequence homology using GET\_HOMOLOGUES software package [110]. We used the core genes from a representative *Xep* genome, Xp91-118, as query to search the remaining 269 genomes using local BLAST [111]. BLAST results were filtered using query coverage and pairwise nucleotide sequence alignment thresholds of 70% each and the sequence was checked for the presence of standard start and stop codons at either end of the gene and gene was removed if both were not present. A total of 887 genes were found to be intact in all 270 genomes. Genes were individually parsed and aligned using MAFFT [112] and concatenated using sequence matrix [113]. The result was a 617.854 Kbp alignment, hereafter referred to as core genes.

The concatenated core gene sequence was used to construct a maximum likelihood (ML) phylogenetic tree using RAxML v.8.2.12 [114]. General time reversible model with gamma distributed rates and invariant sites (GTRGAMMA) was used as the nucleotide substitution model. To account for recombination, the ML tree output from RAxML and concatenated core genome alignment were used as input for ClonalFrameML v1.12 [71].

### Population structure

SNPs were extracted from core genes for hierarchical clustering based on Bayesian analysis of population structure (hierBAPS) algorithm [115], implemented in the ‘hierBAPS’ R package v 1.0.1 [72,116]. For visualization, hierBAPS clusters were added to the phylogenetic tree generated from ClonalFrameML using the ‘ggtree’ package in R [117]. The treemap function in plotly [118] was used to show the relative distribution of clusters across geographic locations. R package ‘ggplot2’ was used to map hierBAPS clusters to countries [119]. The ‘PopGenome’ R package [120] was used to calculate  $F_{ST}$ , Watterson’s theta, nucleotide diversity ( $\pi$ ) [69], and Tajima’s D statistic [70] by geographic location and by hierBAPS cluster.

Assembled genomes were used for calculating average nucleotide identity and pangenome analysis. Average nucleotide identity (ANIb) between strains was calculated using whole genome assemblies with Pyani version 0.2.10 [121]. The pangenome was estimated using Roary v3.12.0 [122] after annotation from Prokka v1.12 [123]. The gene presence absence matrix from Roary (S4 Data) was used as input for generation of NMDS plots using the ‘dplyr’ and ‘ggplot2’ packages from tidyverse [119] and to generate gene accumulation curves for each cluster using package ‘micropan’ [124].

### Bayesian analysis of *X. euvesicatoria* pv. *perforans* divergence times

A whole genome alignment was generated using split k-mer analysis version 2 (SKA2) [125] for all 270 *Xep* strains plus outgroup *X. euvesicatoria* pv. *euvesicatoria* strain 85-10 (NCBI Accession GCA\_000009165.1; S2 Data). The alignment was reduced to variable sites only using Geneious 2023.2.1 (BioMatters Ltd.). A phylogenetic network was calculated from the resulting SNPs using the NeighborNet 2004 algorithm in SplitsTree5 [126,127]. Phylogenetic conflict was indicated between the 259 strains, cluster 10 strains, and outgroup (S1B Fig). Removing the cluster 10 strains did not remove the phylogenetic conflict (reticulations) between *Xep* and *Xee* outgroup. Because the location of the root was not clear, we limited our

dating of the phylogeny of *Xep* to the 259 strains in BAPS clusters 1 through 9 and inferred the root position as part of the analysis. We used Gubbins v. 2.4.1 [73] to remove putative recombinant sites from whole genome alignments generated using SKA2 [125] and the complete genome of Xp91-118 as a reference (GCF\_000192045.2). The resulting alignment was used to infer a phylogenetic tree using the GTRGAMMI model in RAxML version 8.2.10 [128]. The temporal analysis was conducted with BactDating v1.1.1 [74]. The inputs to the BactDating analysis were the maximum likelihood tree and dates of isolation assigned as dates of tips. The rooting of the tree was estimated using the `initRoot` function, which maximizes the correlation between tip date, the year the strain was collected, and root-to-tip branch lengths. Dates of nodes were inferred using the `bactdate` function on the re-rooted tree using a relaxed molecular clock with Markov chain Monte Carlo (MCMC) chains of  $10^6$  iterations. PhyloStems [129] was used to assess the temporal signals within internal clades for interpretation of node date inferences.

We also used BEAST v. 1.10.4 [130] to infer a dated phylogeny. The XML file was manually edited to include the 'ascertained' flag in the alignment block (S3 Data). The HKY nucleotide substitution model with empirical base frequencies and gamma distribution of site-specific rate heterogeneity was used with coalescent Bayesian skyline priors with an uncorrelated relaxed clock for Bayesian phylogenetic inference over MCMC chains of 200 million generations. Adequate mixing was assessed based on a minimum effective sample size of 200 for parameter estimates as calculated by Tracer v. 1.10.4. A maximum clade credibility tree was inferred from the posterior distribution of trees using TreeAnnotator v. 1.10.4, specifying a burn-in of 10% and the 'keep' option for node heights. Trees were visualized in iTOL version 6.9.1 [131].

### Type III effector analysis

A T3E database was generated using amino acid sequences of 63 *Xanthomonas* effectors based on a community-curated list [132] (S6 Table). When available, functional annotations were retrieved from NCBI and Pfam databases [133]. Orthologous sequences were identified with the software BLASTp [134,135], by querying the curated effectors database against the amino acid sequences of the annotated genomes of 270 *Xep* strains. Sequences (BLAST hits) were considered effector orthologs when at or above a threshold of 70 percent identity and 50 percent query coverage. When multiple sequences from the same strain had hits above the thresholds to a particular effector, we used the product of the percent identity and query coverage to select the best overall hit. Sequences with homology to multiple effectors and sequences with evidence of contig breaks were manually removed. Assignment of sequences as effector orthologues was confirmed by performing a clustering analysis of all sequences using the software USEARCH v. 11.0.667 and the algorithm HPC-CLUST [136]. For the duplicated effector XopP, we used a phylogenetic analysis of all sequences to distinguish likely orthologous alleles from the more genetically distant paralogous sequences, which were assigned to XopP2.

Orthologous sequences from each effector were extracted from the annotated genomes, aligned with MAFFT [112], and allelic variants identified [137] to generate a numeric matrix representing presence and allelic variant or absence. Hierarchical clustering analysis of effectors was performed by calculating a distance matrix with function 'dist' with the method 'manhattan', and the function 'hclust' with the method 'complete' from the R package 'vegan' [116,138]. The results were displayed as a heatmap with the package 'gplots' and the function `heatmap.2` [139].

To investigate the presence of positive selection acting on the effector sequences, we used the software HyPhy (Hypothesis Testing using Phylogenies) implementing the methods

FUBAR (Fast, Unconstrained Bayesian Approximation) and MEME (Mixed Effects Model of Evolution) [78, 79]. The Bayesian method FUBAR evaluates pervasive selection, assuming the same rates of synonymous and nonsynonymous substitution per site on all branches. The method MEME uses a maximum likelihood approach to evaluate episodic selection, i.e., selection only a subset of branches of the phylogeny. For each effector gene, a codon-aware alignment was generated with the software PRANK using the codon flag '-c' as settings [140]. RAXML [114] was used to infer a phylogenetic tree with the GTRGAMMA (gamma time-reversible) model of nucleotide substitution. The codon-aware alignment and phylogenetic tree were used as the input files for FUBAR and MEME.

To determine the relationship of the effector profiles with respect to core gene cluster, geographic and temporal distribution, we transformed the dissimilarities in the matrix of effector profiles into distances with non-metric multidimensional scaling (NMDS). We used the Bray-Curtis dissimilarity index, a robust index able to handle missing data that considers the presence and absence of effectors as equally informative, calculated with the package 'vegan' and the function 'metaMDS' [138]. We used a low number of dimensions (K=2) and set try=30 and trymax=500 for random starts to avoid the NMDS getting trapped in local optima. NMDS plots were created with the packages 'ggrepel' and 'ggplot2' [141,142]. Based on the NMDS analysis, we assigned strains to effector clusters, which were plotted on a worldwide map with the packages 'ggplot2' and 'scatterpie' [142,143]. The map was created in R with the packages 'cowplot', 'ggrepel', 'ggspatial', 'libwgeom', 'sf', 'rgeos', 'memisc', 'oz', 'maptools' and 'rnatualearth' with the function 'ne\_countries' [141,144–151]. Geographic coordinates (longitude, latitude) of countries and states (for USA) of collection were obtained with the R package 'googleway' [152] and the function 'mutate\_geocode' from Google maps.

To sequence the putative TAL effector from 2P6S1, native plasmid DNA was isolated using the alkaline lysis method [153]. *EcoRI* digested DNA of the plasmid prep was ligated into vector pLAFR3 [154] restricted with the same enzyme for transformation into *E. coli* DH5α. Clones containing the TAL effector were identified by PCR and analyzed by restriction digest. One clone, designated as p7.1, contained an approx. 5 Kbp *EcoRI* fragment and was selected for Sanger sequencing and phenotype testing. For Sanger sequencing of the TAL repeat region, DNA of p7.1 was restricted with *NsiI* and the internal fragment was ligated into vector pBluescript restricted with *PstI*. Additional pBluescript subclones were made using *BamHI* (~3 Kbp and ~1.1 Kbp) and *BamHI/EcoRI* (~1 Kbp) in order to cover the entire cloned region in p7.1. All clones were transformed into DH5α for sequencing using vector primers T3 and T7.

Copper resistance genes in assembled genomes were identified with BLASTn analysis using *copL* (MBZ2440241.1), *copA* (MBZ2440240.1), and *copB* (MBZ2440239.1) from *Xep* strain Xp2010 as reference sequences [33].

## Supporting information

**S1 Fig. Phylogenetic analysis of 270 *X. euvesicatoria* pv. *perforans* strains.** (A) Maximum likelihood phylogenetic tree of *Xanthomonas euvesicatoria* pv. *perforans* strains based on aligned nucleotide sequences of 887 core genes, also used for Fig 1A. The tree was inferred using RAXML using a GTRGAMMAI substitution model. The tree was rooted using the 11 genetically diverged strains that make up core gene cluster 10. (B) NeighborNet network inferred using SNPs from aligned whole genome sequences, including *X. euvesicatoria* pv. *euvesicatoria* strain 85-10 (bolded) as an outgroup. Core gene cluster 10 strains are highlighted. Reticulations in the network indicate conflicting phylogenetic relationships. (PDF)

**S2 Fig. Accessory genome variation in *Xanthomonas euvesicatoria* pv. *perforans*.** (A) Visualization of pangenome variation by non-metric multidimensional scaling of gene presence-absence for all 270 *X. perforans* strains by BAPS cluster. Ellipses assume a multivariate t-distribution. (C) Increase in gene count with increasing number of strains sampled. Clusters 1 and 2 were represented by the most strains, but other clusters showed similar rates of increase in the pangenome of the cluster. Pangenome matrix used for analysis is available as [S4 Data](#).  
(PDF)

**S3 Fig. Temporal signal in phylogenetic tree of 259 *X. euvesicatoria* pv. *perforans* strains.** (A) Correlation between sampling year and root-to-tip distance in maximum likelihood phylogenetic tree inferred from alignment of whole genome sequences. Output was generated from BactDating R package. (B) Temporal signal within the phylogenetic determined using PhyloStems tool. Nodes with significant temporal signals are indicated with colored circles. Adjusted R-squared values by color are: dark green 0–0.2; light green 0.2–0.4; yellow 0.4–0.6; orange 0.6–0.8; red 0.8–1.  
(PDF)

**S4 Fig. Dated phylogenies of 259 *X. euvesicatoria* pv. *perforans* strains.** (A) Dating using BactDating relaxed clock analysis on RAxML-generated phylogeny. This is the tree shown in Fig 2, shown here without collapsed nodes. (B) Dating of same dataset using BEAST with coalescent Bayesian skyline priors and an uncorrelated relaxed clock.  
(PDF)

**S5 Fig. Frequency of Xop effectors among 270 *Xanthomonas euvesicatoria* pv. *perforans* strains.** The most common allele observed was assigned to allele type 1, second most frequent allele to allele type 2, and so on. Note that alleles classified as pseudogenes included contig breaks, which include assembly errors. For example, all strains appear to have *xopD*, but a repeat caused a contig break in the gene in nearly half of the genomes.  
(PDF)

**S6 Fig. Clustering of 270 *Xanthomonas perforans* effector profiles by non-metric multidimensional scaling and distribution of resulting clusters among geographic regions.** Analysis did not include TAL effectors. (A) The most frequently observed group of effector profiles form cluster A. This cluster of 188 strains is represented as a star in plots B–C, as it is represented in most BAPS core gene clusters (B), most of the sampled tomato production regions (C), and in collections from 1991 to 2017 (D). Clusters were largely defined by low frequency effectors ([S7 Fig](#)). (E) Distribution of strains by effector clusters among sampled countries. Base layer of map is from Natural Earth (<https://www.naturalearthdata.com>).  
(PDF)

**S7 Fig. Variation in type III effector profiles in 270 *Xanthomonas euvesicatoria* pv. *perforans* strains ordered according to NMDS of effector profiles.** Analysis did not include TAL effectors. Type III effectors are in columns and *Xep* strains in rows. Effector status is shown by allele type: absence is indicated by allele type 0 (white), while the most frequent allele observed when the effector is present is allele type 1 (purple), second most frequent is allele type 2 (blue), and so on. Putative pseudogenized effectors are shown as allele 13 (gray). The order of columns was determined by hierarchical clustering analysis, placing similarly distributed effectors adjacent to each other. Order of rows is based on NMDS clustering analysis of effector profiles (see [S6 Fig](#)).  
(PDF)

**S1 Table. Genome data and metadata for *X. euvesicatoria* pv. *perforans* strains.** (A) BAPS cluster assignment for each strain, NCBI Accession for each genome, and associated genome assembly statistics. (B) Additional genomic data used only for ANI comparisons in Fig 1D. (XLSX)

**S2 Table. Genetic diversity statistics by geographic region and BAPS group.** For each geographic region and BAPS group, we calculated: the number of SNPs in the 617854 bp alignment; Watterson's Theta per site; Pi, the average number of differences per site; and Tajima's D. (A) Statistics by country and U.S. state. (B) Statistics by BAPS group. (XLSX)

**S3 Table. Average nucleotide identity (ANIb) comparisons between strains with highly similar core gene sequences collected across continents.** (A) Proportion nucleotide identity. (B) Alignment fraction. (XLSX)

**S4 Table. Putative type III effectors (Xop proteins) found in 270 *X. euvesicatoria* pv. *perforans* assembled genomes.** (A) Summary for each locus. (B) Results by strain. Each different amino acid sequence per gene was assigned a numerical allele type, such that the most common allele observed was assigned to allele type 1. Potential pseudogenes are indicated with "pseudo" and absence indicated with zero. Locus tags refer to JGI IMG annotations (<https://img.jgi.doe.gov>). Reference sequences used for BLAST searches are given in [S5 Table](#). The final column shows the result of BLAST searches for TAL effectors. (XLSX)

**S5 Table. Presence or absence of copper genes (*copLAB*) in 270 *X. euvesicatoria* pv. *perforans* assembled genomes.** Symbols represent gene presence '+' or absence '-'. Contig break in gene is indicated by (+). (XLSX)

**S6 Table. Type III effector database used to query assembled genomes for effector genes.** (XLSX)

**S1 Data. Nucleotide alignment of 887 core genes from 270 *X. euvesicatoria* pv. *perforans* strains.** Alignment is 617,855 bp in FASTA format. (ZIP)

**S2 Data. Nucleotide alignment of variable sites from whole genome alignment of 270 *X. euvesicatoria* pv. *perforans* strains and *X. euvesicatoria* pv. *euvesicatoria* strains 85-10.** (ZIP)

**S3 Data. XML file used for BEAST analysis.** (ZIP)

**S4 Data. Pangenome matrix for 270 *X. euvesicatoria* pv. *perforans* strains.** (ZIP)

## Acknowledgements

We thank D. Ritchie and C. Mauney for providing reference strains and collecting disease samples, respectively. Tina Simonton and Phyllis May provided assistance with strain collection in Ontario. Plant diagnostic laboratories are supported by the USDA-NIFA through the National Plant Diagnostic Network. Training of FIB was supported by Conacyt Mexico and MK by Swedish International Development Cooperation Agency.

## Author contributions

**Conceptualization:** Sujan Timilsina, Fernanda Iruegas-Bocardo, Mustafa O. Jibrin, Jeffrey B. Jones, Gary E. Vallad, Erica M. Goss.

**Data curation:** Sujan Timilsina, Fernanda Iruegas-Bocardo, Mustafa O. Jibrin, Anuj Sharma.

**Formal analysis:** Sujan Timilsina, Fernanda Iruegas-Bocardo, Mustafa O. Jibrin, Anuj Sharma, Aastha Subedi, Amandeep Kaur, Erica M. Goss.

**Funding acquisition:** Pamela D. Roberts, Jeffrey B. Jones, Gary E. Vallad, Erica M. Goss.

**Investigation:** Sujan Timilsina, Fernanda Iruegas-Bocardo, Mustafa O. Jibrin, Aastha Subedi, Gerald V. Minsavage, Jeannie Klein-Gordon, Jeffrey B. Jones, Gary E. Vallad.

**Methodology:** Jose C. Huguet-Tapia.

**Project administration:** Erica M. Goss, Sujan Timilsina.

**Resources:** Mustafa O. Jibrin, Pragya Adhikari, Tika B. Adhikari, Gabriella Cirvilleri, Laura Belen Tapia de la Barrera, Eduardo Bernal, Tom C. Creswell, Tien Thi Kieu Doan, Teresa A. Coutinho, Daniel S. Egel, Rubén Félix-Gastélum, David M. Francis, Misrak Kebede, Melanie Lewis Ivey, Frank J. Louws, Laixin Luo, Elizabeth T. Maynard, Sally A. Miller, Nga Thi Thu Nguyen, Ebrahim Osdaghi, Alice M. Quezado-Duval, Rebecca Roach, Francesca Rotondo, Gail E. Ruhl, Vou M. Shutt, Petcharat Thummabenjapone, Cheryl Trueman.

**Supervision:** Pamela D. Roberts, Jeffrey B. Jones, Gary E. Vallad, Erica M. Goss.

**Visualization:** Sujan Timilsina, Fernanda Iruegas-Bocardo, Anuj Sharma, Erica M. Goss.

**Writing – original draft:** Sujan Timilsina, Fernanda Iruegas-Bocardo, Mustafa O. Jibrin, Jeffrey B. Jones, Gary E. Vallad, Erica M. Goss.

**Writing – review & editing:** Anuj Sharma, Aastha Subedi, Amandeep Kaur, Gerald V. Minsavage, Jose C. Huguet-Tapia, Jeannie Klein-Gordon, Tika B. Adhikari, Gabriella Cirvilleri, Laura Belen Tapia de la Barrera, Eduardo Bernal, Tom C. Creswell, Tien Thi Kieu Doan, Teresa A. Coutinho, Daniel S. Egel, Rubén Félix-Gastélum, David M. Francis, Misrak Kebede, Melanie Lewis Ivey, Frank J. Louws, Laixin Luo, Elizabeth T. Maynard, Sally A. Miller, Nga Thi Thu Nguyen, Ebrahim Osdaghi, Alice M. Quezado-Duval, Rebecca Roach, Francesca Rotondo, Gail E. Ruhl, Vou M. Shutt, Petcharat Thummabenjapone, Cheryl Trueman, Pamela D. Roberts, Jeffrey B. Jones, Gary E. Vallad, Erica M. Goss.

## References

1. Savary S, Willocquet L, Pethybridge SJ, Esker P, McRoberts N, Nelson A. The global burden of pathogens and pests on major food crops. *Nat Ecol Evol*. 2019;3(3):430–9. <https://doi.org/10.1038/s41559-018-0793-y> PMID: [30718852](https://pubmed.ncbi.nlm.nih.gov/30718852/)
2. Ristaino JB, Anderson PK, Bebber DP, Brauman KA, Cunniffe NJ, Fedoroff NV, et al. The persistent threat of emerging plant disease pandemics to global food security. *Proc Natl Acad Sci U S A*. 2021;118(23):e2022239118. <https://doi.org/10.1073/pnas.2022239118> PMID: [34021073](https://pubmed.ncbi.nlm.nih.gov/34021073/)
3. Rizzo DM, Lichtveld M, Mazet JAK, Togami E, Miller SA. Plant health and its effects on food safety and security in a One Health framework: four case studies. *One Health Outlook*. 2021;3:6. <https://doi.org/10.1186/s42522-021-00038-7> PMID: [33829143](https://pubmed.ncbi.nlm.nih.gov/33829143/)
4. Martins PMM, Merfa MV, Takita MA, De Souza AA. Persistence in phytopathogenic bacteria: do we know enough?. *Front Microbiol*. 2018;9:1099. <https://doi.org/10.3389/fmicb.2018.01099> PMID: [29887856](https://pubmed.ncbi.nlm.nih.gov/29887856/)
5. Jeger MJ, Fielder H, Beale T, Szyniszewska AM, Parnell S, Cunniffe NJ. What can be learned by a synoptic review of plant disease epidemics and outbreaks published in 2021?. *Phytopathology*. 2023;113(7):1141–58. <https://doi.org/10.1094/PHYTO-02-23-0069-IA> PMID: [36935375](https://pubmed.ncbi.nlm.nih.gov/36935375/)

6. Mansfield J, Genin S, Magori S, Citovsky V, Sriariyanum M, Ronald P, et al. Top 10 plant pathogenic bacteria in molecular plant pathology. *Mol Plant Pathol*. 2012;13(6):614–29. <https://doi.org/10.1111/j.1364-3703.2012.00804.x> PMID: 22672649
7. Pfeilmeier S, Caly DL, Malone JG. Bacterial pathogenesis of plants: future challenges from a microbial perspective: challenges in bacterial molecular plant pathology. *Mol Plant Pathol*. 2016;17(8):1298–313. <https://doi.org/10.1111/mpp.12427> PMID: 27170435
8. Morris CE, Sands DC, Vinatzer BA, Glaux C, Guilbaud C, Buffière A, et al. The life history of the plant pathogen *Pseudomonas syringae* is linked to the water cycle. *ISME J*. 2008;2(3):321–34. <https://doi.org/10.1038/ismej.2007.113> PMID: 18185595
9. Roman-Reyna V, Sharma A, Toth H, Konkelt Z, Omiotek N, Murthy S, et al. Live tracking of a plant pathogen outbreak reveals rapid and successive, multidecade plasmid reduction. *mSystems*. 2024;9(2):e0079523. <https://doi.org/10.1128/msystems.00795-23> PMID: 38275768
10. Campos PE, Groot Crego C, Boyer K, Gaudeul M, Baider C, Richard D, et al. First historical genome of a crop bacterial pathogen from herbarium specimen: insights into citrus canker emergence. *PLoS Pathog*. 2021;17(7):e1009714. <https://doi.org/10.1371/journal.ppat.1009714> PMID: 34324594
11. Xu J, Zhang Y, Li J, Teper D, Sun X, Jones D, et al. Phylogenomic analysis of 343 *Xanthomonas citri* pv. *citri* strains unravels introduction history and dispersal paths. *PLoS Pathog*. 2023;19(12):e1011876. <https://doi.org/10.1371/journal.ppat.1011876> PMID: 38100539
12. Weisberg AJ, Davis EW 2nd, Tabima J, Belcher MS, Miller M, Kuo C-H, et al. Unexpected conservation and global transmission of agrobacterial virulence plasmids. *Science*. 2020;368(6495):eaba5256. <https://doi.org/10.1126/science.aba5256> PMID: 32499412
13. Castillo AI, Chacón-Díaz C, Rodríguez-Murillo N, Coletta-Filho HD, Almeida RPP. Impacts of local population history and ecology on the evolution of a globally dispersed pathogen. *BMC Genomics*. 2020;21(1):369. <https://doi.org/10.1186/s12864-020-06778-6> PMID: 32434538
14. Castillo AI, Bojanini I, Chen H, Kandel PP, De La Fuente L, Almeida RPP. Allopatric plant pathogen population divergence following disease emergence. *Appl Environ Microbiol*. 2021;87(7):e02095-20. <https://doi.org/10.1128/AEM.02095-20> PMID: 33483307
15. Jibrin MO, Sharma A, Mavian CN, Timilsina S, Kaur A, Iruegas-Bocardo F, et al. Phylodynamic Insights into Global Emergence and Diversification of the Tomato Pathogen *Xanthomonas hortorum* pv. *gardneri*. *Mol Plant Microbe Interact*. 2024;37(10):712–20. <https://doi.org/10.1094/MPMI-04-24-0035-R> PMID: 38949619
16. Subedi A, Barrera LBT de la, Ivey ML, Egel DS, Kebede M, Kara S, et al. Population genomics reveals an emerging lineage of *xanthomonas perforans* on pepper. *Phytopathology*. 2024;114(1):241–50. <https://doi.org/10.1094/PHYTO-04-23-0128-R> PMID: 37432099
17. Dupas E, Durand K, Rieux A, Briand M, Pruvost O, Cuntly A, et al. Suspicions of two bridgehead invasions of *Xylella fastidiosa* subsp. *multiplex* in France. *Commun Biol*. 2023;6(1):103. <https://doi.org/10.1038/s42003-023-04499-6> PMID: 36707697
18. Constantin EC, Cleenwerck I, Maes M, Baeyen S, Van Malderghem C, De Vos P, et al. Genetic characterization of strains named as *Xanthomonas axonopodis* pv. *dieffenbachiae* leads to a taxonomic revision of the *X. axonopodis* species complex. *Plant Pathology*. 2015;65(5):792–806. <https://doi.org/10.1111/ppa.12461>
19. Jones JB. A Third Tomato Race of *Xanthomonas campestris* pv. *vesicatoria*. *Plant Dis*. 1995;79(4):395. <https://doi.org/10.1094/pd-79-0395>
20. Jones JB, Lacy GH, Bouzar H, Stall RE, Schaad NW. Reclassification of the *xanthomonads* associated with bacterial spot disease of tomato and pepper. *Syst Appl Microbiol*. 2004;27(6):755–62. <https://doi.org/10.1078/0723202042369884> PMID: 15612634
21. Potnis N, Timilsina S, Strayer A, Shantharaj D, Barak JD, Paret ML, et al. Bacterial spot of tomato and pepper: diverse *Xanthomonas* species with a wide variety of virulence factors posing a worldwide challenge. *Mol Plant Pathol*. 2015;16(9):907–20. <https://doi.org/10.1111/mpp.12244> PMID: 25649754
22. Osdaghi E, Jones JB, Sharma A, Goss EM, Abrahamian P, Newberry EA, et al. A centenary for bacterial spot of tomato and pepper. *Mol Plant Pathol*. 2021;22(12):1500–19. <https://doi.org/10.1111/mpp.13125> PMID: 34472193
23. Potnis N. Harnessing eco-evolutionary dynamics of *xanthomonads* on tomato and pepper to tackle new problems of an old disease. *Annu Rev Phytopathol*. 2021;59(1):289–310. <https://doi.org/10.1146/annurev-phyto-020620-101612>
24. Jibrin MO, Timilsina S, Minsavage GV, Vallad GE, Roberts PD, Goss EM, et al. Bacterial Spot of Tomato and Pepper in Africa: Diversity, Emergence of T5 Race, and Management. *Front Microbiol*. 2022;13:835647. <https://doi.org/10.3389/fmicb.2022.835647> PMID: 35509307

25. Abbasi PA, Khabbaz SE, Weselowski B, Zhang L. Occurrence of copper-resistant strains and a shift in *Xanthomonas* spp. causing tomato bacterial spot in Ontario. *Can J Microbiol*. 2015;61(10):753–61. <https://doi.org/10.1139/cjm-2015-0228> PMID: 26308592
26. Klein-Gordon JM, Xing Y, Garrett KA, Abrahamian P, Paret ML, Minsavage GV, et al. Assessing changes and associations in the *xanthomonas perforans* population across florida commercial tomato fields via a statewide survey. *Phytopathology*. 2021;111(6):1029–41. <https://doi.org/10.1094/PHYTO-09-20-0402-R> PMID: 33048630
27. Lewis Ivey ML, Strayer A, Sidhu JK, Minsavage GV. Bacterial leaf spot of tomato (*solanum lycopersicum*) in louisiana is caused by *xanthomonas perforans*, tomato race 4. *Plant Disease*. 2016;100(6):1233–1233. <https://doi.org/10.1094/pdis-12-15-1451-pdn>
28. Egel DS, Jones JB, Minsavage GV, Creswell T, Ruhl G, Maynard E, et al. Distribution and characterization of *xanthomonas* strains causing bacterial spot of tomato in Indiana. *Plant Health Progress*. 2018;19(4):319–21. <https://doi.org/10.1094/php-07-18-0041-br>
29. Abrahamian P, Jones JB, Vallad GE. Efficacy of copper and copper alternatives for management of bacterial spot on tomato under transplant and field production. *Crop Protection*. 2019;126:104919. <https://doi.org/10.1016/j.cropro.2019.104919>
30. Adhikari P, Adhikari TB, Timilsina S, Meadows I, Jones JB, Panthee DR, et al. Phenotypic and genetic diversity of *xanthomonas perforans* populations from tomato in North Carolina. *Phytopathology*. 2019;109(9):1533–43. <https://doi.org/10.1094/PHYTO-01-19-0019-R> PMID: 31038016
31. Rotondo F, Bernal E, Ma X, Ivey MLL, Sahin F, Francis DM, et al. Shifts in *Xanthomonas* spp. causing bacterial spot in processing tomato in the Midwest of the United States. *Canad Plant Pathol*. 2022;44(5):652–67. <https://doi.org/10.1080/07060661.2022.2047788>
32. Khanal S, Hind SR, Babadoost M. Occurrence of copper-resistant *xanthomonas perforans* and *x. gardneri* in illinois tomato Fields. *Plant Health Progress*. 2020;21(4):338–44. <https://doi.org/10.1094/php-06-20-0048-rs>
33. Bibi S, Weis K, Kaur A, Bhandari R, Goss E, Jones JB, et al. A brief evaluation of a copper resistance mobile genetic island in the bacterial leaf spot pathogen *xanthomonas euvesicatoria* pv. *perforans*. *Phytopathology*. 2023;113(8):1394–8. <https://doi.org/10.1094/PHYTO-02-23-0077-SC> PMID: 37097444
34. Roach R, Mann R, Gambley CG, Shivas RG, Rodoni B. Identification of *Xanthomonas* species associated with bacterial leaf spot of tomato, capsicum and chilli crops in eastern Australia. *Eur J Plant Pathol*. 2017;150(3):595–608. <https://doi.org/10.1007/s10658-017-1303-9>
35. Burlakoti RR, Hsu C-F, Chen J-R, Wang J-F. Population dynamics of *xanthomonads* associated with bacterial spot of tomato and pepper during 27 years across Taiwan. *Plant Dis*. 2018;102(7):1348–56. <https://doi.org/10.1094/PDIS-04-17-0465-RE> PMID: 30673574
36. Kebede M, Timilsina S, Ayalew A, Admassu B, Potnis N, Minsavage GV, et al. Molecular characterization of *Xanthomonas* strains responsible for bacterial spot of tomato in Ethiopia. *Eur J Plant Pathol*. 2014;140(4):677–88. <https://doi.org/10.1007/s10658-014-0497-3>
37. Timilsina S, Jibrin MO, Potnis N, Minsavage GV, Kebede M, Schwartz A, et al. Multilocus sequence analysis of *xanthomonads* causing bacterial spot of tomato and pepper plants reveals strains generated by recombination among species and recent global spread of *Xanthomonas gardneri*. *Appl Environ Microbiol*. 2015;81(4):1520–9. <https://doi.org/10.1128/AEM.03000-14> PMID: 25527544
38. Hamza AA, Robène-Soustrade I, Jouen E, Gagnevin L, Lefeuvre P, Chiroleu F, et al. Genetic and pathological diversity among *xanthomonas* strains responsible for bacterial spot on tomato and pepper in the Southwest Indian Ocean Region. *Plant Dis*. 2010;94(8):993–9. <https://doi.org/10.1094/PDIS-94-8-0993> PMID: 30743480
39. Araújo ER, Costa JR, Ferreira MASV, Quezado-Duval AM. Widespread distribution of *Xanthomonas perforans* and limited presence of *X. gardneri* in Brazil. *Plant Pathol*. 2017;66(1):159–68. doi: <https://doi.org/10.1111/ppa.12543>
40. Osdaghi E, Taghavi SM, Hamzehzarghani H, Lamichhane JR. Occurrence and characterization of the bacterial spot pathogen *xanthomonas euvesicatoria* on pepper in Iran. *J Phytopathol*. 2016;164(10):722–34. <https://doi.org/10.1111/jph.12493>
41. Osdaghi E, Taghavi SM, Hamzehzarghani H, Fazliarab A, Lamichhane JR. Monitoring the occurrence of tomato bacterial spot and range of the causal agent *Xanthomonas perforans* in Iran. *Plant Pathol*. 2016;66(6):990–1002. <https://doi.org/10.1111/ppa.12642>
42. Osdaghi E, Taghavi SM, Koebnik R, Lamichhane JR. Multilocus sequence analysis reveals a novel phylogroup of *Xanthomonas euvesicatoria* pv. *perforans* causing bacterial spot of tomato in Iran. *Plant Pathol*. 2018;67(7):1601–11. <https://doi.org/10.1111/ppa.12864>

43. Abrahamian P, Timilsina S, Minsavage GV, Potnis N, Jones JB, Goss EM, et al. Molecular Epidemiology of xanthomonas perforans outbreaks in tomato plants from transplant to field as determined by single-nucleotide polymorphism analysis. *Appl Environ Microbiol*. 2019;85(18):e01220-19. <https://doi.org/10.1128/AEM.01220-19> PMID: 31253682
44. Modor Intelligence. Tomato Seeds Market - Growth, Trends, COVID-10 Impact, and Forecasts (2022 - 2027) 2022. Available from: Available at: <https://www.mordorintelligence.com/industry-reports/tomato-seeds-market> [Accessed August 31, 2022].
45. Abrahamian P, Sharma A, Jones JB, Vallad GE. Dynamics and spread of bacterial spot epidemics in tomato transplants grown for field production. *Plant Dis*. 2021;105(3):566–75. <https://doi.org/10.1094/PDIS-05-20-0945-RE> PMID: 32865478
46. Hert AP, Roberts PD, Momol MT, Minsavage GV, Tudor-Nelson SM, Jones JB. Relative importance of bacteriocin-like genes in antagonism of *Xanthomonas perforans* tomato race 3 to *Xanthomonas euvesicatoria* tomato race 1 strains. *Appl Environ Microbiol*. 2005;71(7):3581–8. <https://doi.org/10.1128/AEM.71.7.3581-3588.2005> PMID: 16000765
47. Timilsina S, Pereira-Martin JA, Minsavage GV, Iruegas-Bocardo F, Abrahamian P, Potnis N, et al. Multiple recombination events drive the current genetic structure of xanthomonas perforans in Florida. *Front Microbiol*. 2019;10:448. <https://doi.org/10.3389/fmicb.2019.00448> PMID: 30930868
48. Jibrin MO, Potnis N, Timilsina S, Minsavage GV, Vallad GE, Roberts PD, et al. Genomic inference of recombination-mediated evolution in xanthomonas euvesicatoria and X. perforans. *Appl Environ Microbiol*. 2018;84(13):e00136-18. <https://doi.org/10.1128/AEM.00136-18> PMID: 29678917
49. Abrahamian P, Timilsina S, Minsavage GV, Kc S, Goss EM, Jones JB, et al. The type iii effector avrbst enhances xanthomonas perforans fitness in field-grown Tomato. *Phytopathology*. 2018;108(12):1355–62. <https://doi.org/10.1094/PHYTO-02-18-0052-R> PMID: 29905507
50. Newberry EA, Bhandari R, Minsavage GV, Timilsina S, Jibrin MO, Kemble J, et al. Independent evolution with the gene flux originating from multiple xanthomonas species explains genomic heterogeneity in *Xanthomonas perforans*. *Appl Environ Microbiol*. 2019;85(20):e00885-19. <https://doi.org/10.1128/AEM.00885-19> PMID: 31375496
51. Sharma A, Timilsina S, Abrahamian P, Minsavage GV, Colee J, Ojiambo PS, et al. Need for speed: bacterial effector XopJ2 is associated with increased dispersal velocity of *Xanthomonas perforans*. *Environ Microbiol*. 2021;23(10):5850–65. <https://doi.org/10.1111/1462-2920.15541> PMID: 33891376
52. Klein-Gordon JM, Guingab-Cagmat J, Minsavage GV, Meke L, Vallad GE, Goss EM, et al. Strength in numbers: density-dependent volatile-induced antimicrobial activity by *Xanthomonas perforans*. *Phytopathology*. 2023;113(2):160–9. <https://doi.org/10.1094/PHYTO-04-22-0131-R> PMID: 36129764
53. Bernal E, Rotondo F, Roman-Reyna V, Klass T, Timilsina S, Minsavage GV, et al. Migration drives the replacement of xanthomonas perforans races in the absence of widely deployed resistance. *Front Microbiol*. 2022;13:826386. <https://doi.org/10.3389/fmicb.2022.826386> PMID: 35369455
54. Klein-Gordon JM, Timilsina S, Xing Y, Abrahamian P, Garrett KA, Jones JB, et al. Whole genome sequences reveal the *Xanthomonas perforans* population is shaped by the tomato production system. *ISME J*. 2022;16(2):591–601. <https://doi.org/10.1038/s41396-021-01104-8> PMID: 34489540
55. Richard D, Boyer C, Lefeuvre P, Canteros BI, Beni-Madhu S, Portier P, et al. Complete genome sequences of six copper-resistant xanthomonas strains causing bacterial spot of solanaceous plants, belonging to *X. gardneri*, *X. euvesicatoria*, and *X. vesicatoria*, using long-read technology. *Genome Announc*. 2017;5(8):e01693-16. <https://doi.org/10.1128/genomeA.01693-16> PMID: 28232425
56. Timilsina S, Potnis N, Newberry EA, Liyanapathirana P, Iruegas-Bocardo F, White FF, et al. *Xanthomonas* diversity, virulence and plant-pathogen interactions. *Nat Rev Microbiol*. 2020;18(8):415–27. <https://doi.org/10.1038/s41579-020-0361-8> PMID: 32346148
57. Roach R, Mann R, Gambley CG, Chapman T, Shivas RG, Rodoni B. Genomic sequence analysis reveals diversity of Australian *Xanthomonas* species associated with bacterial leaf spot of tomato, capsicum and chilli. *BMC Genomics*. 2019;20(1):310. <https://doi.org/10.1186/s12864-019-5600-x> PMID: 31014247
58. Schwartz AR, Potnis N, Timilsina S, Wilson M, Patané J, Martins J Jr, et al. Phylogenomics of *Xanthomonas* field strains infecting pepper and tomato reveals diversity in effector repertoires and identifies determinants of host specificity. *Front Microbiol*. 2015;6:535. <https://doi.org/10.3389/fmicb.2015.00535> PMID: 26089818
59. Horvath DM, Stall RE, Jones JB, Pauly MH, Vallad GE, Dahlbeck D, et al. Transgenic resistance confers effective field level control of bacterial spot disease in tomato. *PLoS One*. 2012;7(8):e42036. <https://doi.org/10.1371/journal.pone.0042036> PMID: 22870280

60. Subedi A, Minsavage GV, Jones JB, Goss EM, Roberts PD. Exploring diversity of bacterial spot associated xanthomonas population of pepper in Southwest Florida. *Plant Dis.* 2023;107(10):2978–85. <https://doi.org/10.1094/PDIS-10-22-2484-RE> PMID: 36856653
61. Cesbron S, Briand M, Essakhi S, Gironde S, Boureau T, Manceau C, et al. Comparative genomics of pathogenic and nonpathogenic strains of xanthomonas arboricola unveil molecular and evolutionary events linked to pathoadaptation. *Front Plant Sci.* 2015;6:1126. <https://doi.org/10.3389/fpls.2015.01126> PMID: 26734033
62. McCann HC. Skirmish or war: the emergence of agricultural plant pathogens. *Curr Opin Plant Biol.* 2020;56:147–52. <https://doi.org/10.1016/j.pbi.2020.06.003> PMID: 32712539
63. Strayer AL, Jeyaprakash A, Minsavage GV, Timilsina S, Vallad GE, Jones JB, et al. A Multiplex real-time pcr assay differentiates four xanthomonas species associated with bacterial spot of tomato. *Plant Dis.* 2016;100(8):1660–8. <https://doi.org/10.1094/PDIS-09-15-1085-RE> PMID: 30686244
64. Aiello D, Scuderi G, Vitale A, Firrao G, Polizzi G, Cirvilleri G. A pith necrosis caused by Xanthomonas perforans on tomato plants. *Eur J Plant Pathol.* 2013;137(1):29–41. <https://doi.org/10.1007/s10658-013-0214-7>
65. Jibrin MO, Timilsina S, Potnis N, Minsavage GV, Shenge KC, Akpa AD, et al. First report of atypical xanthomonas euvesicatoria strains causing bacterial spot of Tomato in Nigeria. *Plant Dis.* 2015;99(3):415. <https://doi.org/10.1094/PDIS-09-14-0952-PDN> PMID: 30699707
66. Timilsina S, Abrahamian P, Potnis N, Minsavage GV, White FF, Staskawicz BJ, et al. Analysis of sequenced genomes of xanthomonas perforans identifies candidate targets for resistance breeding in Tomato. *Phytopathology.* 2016;106(10):1097–104. <https://doi.org/10.1094/PHYTO-03-16-0119-FI> PMID: 27392180
67. Potnis N, Krasileva K, Chow V, Almeida NF, Patil PB, Ryan RP, et al. Comparative genomics reveals diversity among xanthomonads infecting tomato and pepper. *BMC Genomics.* 2011;12:146. <https://doi.org/10.1186/1471-2164-12-146> PMID: 21396108
68. Hudson RR, Slatkin M, Maddison WP. Estimation of levels of gene flow from DNA sequence data. *Genetics.* 1992;132(2):583–9. <https://doi.org/10.1093/genetics/132.2.583> PMID: 1427045
69. Nei M. Molecular evolutionary genetics. New York: Columbia Univ. Press; 1987.
70. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics.* 1989;123(3):585–95. <https://doi.org/10.1093/genetics/123.3.585> PMID: 2513255
71. Didelot X, Wilson DJ. ClonalFrameML: efficient inference of recombination in whole bacterial genomes. *PLoS Comput Biol.* 2015;11(2):e1004041. <https://doi.org/10.1371/journal.pcbi.1004041> PMID: 25675341
72. Tonkin-Hill G, Lees JA, Bentley SD, Frost SDW, Corander J. RhierBAPS: an R implementation of the population clustering algorithm hierBAPS. *Wellcome Open Res.* 2018;3:93. <https://doi.org/10.12688/wellcomeopenres.14694.1> PMID: 30345380
73. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* 2015;43(3):e15. <https://doi.org/10.1093/nar/gku1196> PMID: 25414349
74. Didelot X, Croucher NJ, Bentley SD, Harris SR, Wilson DJ. Bayesian inference of ancestral dates on bacterial phylogenetic trees. *Nucleic Acids Res.* 2018;46(22):e134. <https://doi.org/10.1093/nar/gky783> PMID: 30184106
75. Stall RE, Jones JB, Minsavage GV. Durability of resistance in tomato and pepper to xanthomonads causing bacterial spot. *Annu Rev Phytopathol.* 2009;47:265–84. <https://doi.org/10.1146/annurev-phyto-080508-081752> PMID: 19400644
76. Sharlach M, Dahlbeck D, Liu L, Chiu J, Jiménez-Gómez JM, Kimura S, et al. Fine genetic mapping of RXopJ4, a bacterial spot disease resistance locus from Solanum pennellii LA716. *Theor Appl Genet.* 2012;126(3):601–9. <https://doi.org/10.1007/s00122-012-2004-6>
77. Sharma A, Iruegas-Bocardo F, Bibi S, Chen Y-C, Kim J-G, Abrahamian P, et al. Multiple acquisitions of xopj2 effectors in populations of xanthomonas perforans. *Mol Plant Microbe Interact.* 2024;37(10):736–47. <https://doi.org/10.1094/MPMI-05-24-0048-R> PMID: 39102648
78. Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, et al. FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol Biol Evol.* 2013;30(5):1196–205. <https://doi.org/10.1093/molbev/mst030> PMID: 23420840
79. Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL. Detecting individual sites subject to episodic diversifying selection. *PLoS Genet.* 2012;8(7):e1002764. <https://doi.org/10.1371/journal.pgen.1002764> PMID: 22807683

80. Behlau F, Canteros BI, Minsavage GV, Jones JB, Graham JH. Molecular characterization of copper resistance genes from *Xanthomonas citri* subsp. *citri* and *Xanthomonas alfalfae* subsp. *citrumelonis*. *Appl Environ Microbiol*. 2011;77(12):4089–96. <https://doi.org/10.1128/AEM.03043-10> PMID: 21515725
81. Behlau F, Gochez AM, Lugo AJ, Elibox W, Minsavage GV, Potnis N, et al. Characterization of a unique copper resistance gene cluster in *Xanthomonas campestris* pv. *campestris* isolated in Trinidad, West Indies. *Eur J Plant Pathol*. 2016;147(3):671–81. <https://doi.org/10.1007/s10658-016-1035-2>
82. Stall RE. Linkage of copper resistance and avirulence loci on a self-transmissible plasmid in *Xanthomonas campestris* pv. *vesicatoria*. *Phytopathology*. 1986;76(2):240. <https://doi.org/10.1094/phyto-76-240>
83. Gitaitis R, Walcott R. The epidemiology and management of seedborne bacterial diseases. *Annu Rev Phytopathol*. 2007;45:371–97. <https://doi.org/10.1146/annurev.phyto.45.062806.094321> PMID: 17474875
84. Munkvold GP. Seed pathology progress in academia and industry. *Annu Rev Phytopathol*. 2009;47:285–311. <https://doi.org/10.1146/annurev-phyto-080508-081916> PMID: 19400648
85. Bank W. United States Seed; vegetable seed, of a kind used for sowing imports by country in 2019. Available from: <https://wits.worldbank.org/trade/comtrade/en/country/USA/year/2019/tradeflow/Imports/partner/ALL/product/1209912019> [Accessed 2022-09-15].
86. Jones JB. Survival of *Xanthomonas campestris* pv. *vesicatoria* in Florida on Tomato Crop Residue, Weeds, Seeds, and Volunteer Tomato Plants. *Phytopathology*. 1986;76(4):430. <https://doi.org/10.1094/phyto-76-430>
87. Jones JB, Bouzar H, Stall RE, Almira EC, Roberts PD, Bowen BW, et al. Systematic analysis of xanthomonads (*Xanthomonas* spp.) associated with pepper and tomato lesions. *Int J Syst Evol Microbiol*. 2000;50 Pt 3:1211–9. <https://doi.org/10.1099/00207713-50-3-1211> PMID: 10843065
88. Dutta B, Gitaitis R, Sanders H, Booth C, Smith S, Langston DB Jr. Role of blossom colonization in pepper seed infestation by *Xanthomonas euvesicatoria*. *Phytopathology*. 2014;104(3):232–9. <https://doi.org/10.1094/PHYTO-05-13-0138-R> PMID: 24111576
89. Giovanardi D, Biondi E, Ignjatov M, Jevtić R, Stefani E. Impact of bacterial spot outbreaks on the phytosanitary quality of tomato and pepper seeds. *Plant Pathology*. 2018;67(5):1168–76. <https://doi.org/10.1111/ppa.12839>
90. Parajuli A, Subedi A, Timilsina S, Minsavage GV, Kenyon L, Chen J-R, et al. Phenotypic and genetic diversity of xanthomonads isolated from pepper (*Capsicum* spp.) in Taiwan from 1989 to 2019. *Phytopathology*. 2024;114(9):2033–44. <https://doi.org/10.1094/PHYTO-11-23-0449-R> PMID: 38809758
91. Iruegas-Bocardo F, Weisberg AJ, Riutta ER, Kilday K, Bonkowski JC, Creswell T, et al. Whole genome sequencing-based tracing of a 2022 introduction and outbreak of *xanthomonas hortorum* pv. *pelargonii*. *Phytopathology*. 2023;113(6):975–84. <https://doi.org/10.1094/PHYTO-09-22-0321-R> PMID: 36515656
92. Emsweller S. Fundamental procedures in breeding crops. *The Yearbook of Agriculture U.S. Department of Agriculture*; 1961:127–33.
93. Dorst J. Een en twintigste beschrijvende rassenlijst voor landbouwgewassen. Wageningen, Rijkscommissie voor de samenstelling van de rassenlijst voor landbouwgewassen; 1946. p. 221.
94. Gedgaew C, Simaraks S, Rambo AT. Trends in hybrid tomato seed production under contract farming in Northeast Thailand. *Southeast Asian Studies*. 2017;6(2):339–3.
95. Known-You Nursery & Seed Production Co. Ltd. Production and marketing of F1 hybrid tomato seed in Taiwan. Ed. R. Cowell, 1st International Symposium on Tropical Tomato; 1978.
96. Kloppenburg JR. First the seed: The political economy of plant biotechnology, 1492–2000. (2nd edition). Madison, WI: University of Wisconsin Press; 2004.
97. Kaur A, Minsavage GV, Potnis N, Jones JB, Goss EM. Evolution of copper resistance in *Xanthomonas euvesicatoria* pv. *perforans* population. *mSystems*. 2024;9(12):e0142724. <https://doi.org/10.1128/msystems.01427-24> PMID: 39584814
98. Quibod IL, Atieza-Grande G, Oreiro EG, Palmos D, Nguyen MH, Coronejo ST, et al. The Green Revolution shaped the population structure of the rice pathogen *Xanthomonas oryzae* pv. *oryzae*. *ISME J*. 2020;14(2):492–505. <https://doi.org/10.1038/s41396-019-0545-2> PMID: 31666657
99. McDonald BA, Linde C. Pathogen population genetics, evolutionary potential, and durable resistance. *Annu Rev Phytopathol*. 2002;40:349–79. <https://doi.org/10.1146/annurev.phyto.40.120501.101443> PMID: 12147764

100. Gassmann W, Dahlbeck D, Chesnokova O, Minsavage GV, Jones JB, Staskawicz BJ. Molecular evolution of virulence in natural field strains of *Xanthomonas campestris* pv. *vesicatoria*. *J Bacteriol*. 2000;182(24):7053–9. <https://doi.org/10.1128/JB.182.24.7053-7059.2000> PMID: [11092868](#)
101. Haefner BJ, McCrudden TH, Martin GB. Ptr1 is a CC-NLR immune receptor that mediates recognition of diverse bacterial effectors in multiple solanaceous plants. *Physiol Mole Plant Pathol*. 2023;125:101997. <https://doi.org/10.1016/j.pmpp.2023.101997>
102. Ortega A, Seong K, Schultink A, de Toledo Thomazella DP, Seo E, Zhang E, et al. CRISPR/Cas9-mediated editing of Bs5 and Bs5L in tomato leads to resistance against *Xanthomonas*. *Plant Biotechnol J*. 2024;22(10):2785–7. <https://doi.org/10.1111/pbi.14404> PMID: [39001586](#)
103. Torelli E, Aiello D, Polizzi G, Firrao G, Cirvilleri G. Draft genome of a *Xanthomonas perforans* strain associated with pith necrosis. *FEMS Microbiol Lett*. 2015;362(4):10.1093/femsle/fnv001. <https://doi.org/10.1093/femsle/fnv001> PMID: [25724775](#)
104. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet j*. 2011;17(1):10. <https://doi.org/10.14806/ej.17.1.200>
105. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 2012;19(5):455–77. <https://doi.org/10.1089/cmb.2012.0021> PMID: [22506599](#)
106. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357–9. <https://doi.org/10.1038/nmeth.1923> PMID: [22388286](#)
107. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One*. 2014;9(11):e112963. <https://doi.org/10.1371/journal.pone.0112963> PMID: [25409509](#)
108. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res*. 2015;25(7):1043–55. <https://doi.org/10.1101/gr.186072.114> PMID: [25977477](#)
109. Markowitz VM, Chen I-MA, Palaniappan K, Chu K, Szeto E, Grechkin Y, et al. IMG: the Integrated Microbial Genomes database and comparative analysis system. *Nucleic Acids Res*. 2012;40(Database issue):D115–22. <https://doi.org/10.1093/nar/gkr1044> PMID: [22194640](#)
110. Contreras-Moreira B, Vinuesa P. GET\_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. *Appl Environ Microbiol*. 2013;79(24):7696–701. <https://doi.org/10.1128/AEM.02411-13> PMID: [24096415](#)
111. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215(3):403–10. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2) PMID: [2231712](#)
112. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013;30(4):772–80. <https://doi.org/10.1093/molbev/mst010> PMID: [23329690](#)
113. Vaidya G, Lohman DJ, Meier R. SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics*. 2011;27(2):171–80. <https://doi.org/10.1111/j.1096-0031.2010.00329.x> PMID: [34875773](#)
114. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014;30(9):1312–3. <https://doi.org/10.1093/bioinformatics/btu033> PMID: [24451623](#)
115. Cheng L, Connor TR, Sirén J, Aanensen DM, Corander J. Hierarchical and spatially explicit clustering of DNA sequences with BAPS software. *Mol Biol Evol*. 2013;30(5):1224–8. <https://doi.org/10.1093/molbev/mst028> PMID: [23408797](#)
116. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>
117. Yu G. Using ggtree to Visualize Data on Tree-Like Structures. *Curr Protoc Bioinformatics*. 2020;69(1):e96. <https://doi.org/10.1002/cpbi.96> PMID: [32162851](#)
118. Sievert C. Interactive web-based data visualization with R, plotly, and shiny. Florida: CRC Press; 2020.
119. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, et al. Welcome to the Tidyverse. *JOSS*. 2019;4(43):1686. <https://doi.org/10.21105/joss.01686>
120. Pfeifer B, Wittelsbürger U, Ramos-Onsins SE, Lercher MJ. PopGenome: an efficient Swiss army knife for population genomic analyses in R. *Mol Biol Evol*. 2014;31(7):1929–36. <https://doi.org/10.1093/molbev/msu136> PMID: [24739305](#)

121. Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Anal Methods*. 2016;8(1):12–24. <https://doi.org/10.1039/c5ay02550h>
122. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, et al. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics*. 2015;31(22):3691–3. <https://doi.org/10.1093/bioinformatics/btv421> PMID: 26198102
123. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30(14):2068–9. <https://doi.org/10.1093/bioinformatics/btu153> PMID: 24642063
124. Snipen L, Liland KH. micropan: an R-package for microbial pan-genomics. *BMC Bioinformatics*. 2015;16:79. <https://doi.org/10.1186/s12859-015-0517-0> PMID: 25888166
125. Harris SR. SKA: split kmer analysis toolkit for bacterial genomic epidemiology. 2018. <https://doi.org/10.1101/453142>
126. Huson DH. SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics*. 1998;14(1):68–73. <https://doi.org/10.1093/bioinformatics/14.1.68> PMID: 9520503
127. Huson DH, Bryant D. Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol*. 2006;23(2):254–67. <https://doi.org/10.1093/molbev/msj030> PMID: 16221896
128. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*. 2006;22(21):2688–90. <https://doi.org/10.1093/bioinformatics/btl446> PMID: 16928733
129. Doizy A, Prin A, Cornu G, Chiroleu F, Rieux A. Phylostems: a new graphical tool to investigate temporal signal of heterochronous sequences datasets. *Bioinform Adv*. 2023;3(1):vbad026. <https://doi.org/10.1093/bioadv/vbad026> PMID: 36936370
130. Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol*. 2018;4(1):vey016. <https://doi.org/10.1093/ve/vey016> PMID: 29942656
131. Letunic I, Bork P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res*. 2024;52(W1):W78–82. <https://doi.org/10.1093/nar/gkae268> PMID: 38613393
132. Costa J, Pothier JF, Bosis E, Boch J, Kölliker R, Koebnik R. A Community-curated dokuwiki resource on diagnostics, diversity, pathogenicity, and genetic control of xanthomonads. *Mol Plant Microbe Interact*. 2024;37(3):347–53. <https://doi.org/10.1094/MPMI-11-23-0184-FI> PMID: 38114082
133. Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, et al. The Pfam protein families database. *Nucleic Acids Res*. 2010;38(Database issue):D211–22. <https://doi.org/10.1093/nar/gkp985> PMID: 19920124
134. Boratyn GM, Camacho C, Cooper PS, Coulouris G, Fong A, Ma N, et al. BLAST: a more efficient report with usability improvements. *Nucleic Acids Res*. 2013;41(Web Server issue):W29–33. <https://doi.org/10.1093/nar/gkt282> PMID: 23609542
135. Altschul SF. BLAST Algorithm. *ELSci*. 2014. <https://doi.org/10.1002/9780470015902.a0005253.pub2>
136. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*. 2010;26(19):2460–1. <https://doi.org/10.1093/bioinformatics/btq461> PMID: 20709691
137. Takebayashi N. uniqueHaplo.pl, <http://raven.wrrb.uaf.edu/~ntakebay/teaching/programming/perl-scripts/uniqHaplo.pl>, accessed 4-01-23. 2015.
138. Oksanen J, Simpson GL, Blanchet FG, Kindt R, Legendre P, Minchin PR, et al. vegan: Community Ecology Package. R package version 2.5-6. <https://CRAN.R-project.org/package=vegan>.
139. Warnes G, Bolker B, Bonebakker L, Gentleman R, Huber W, Liaw A, et al. gplots: Various R Programming Tools for Plotting Data. R package version 3.1.3, >. 2020. <https://CRAN.R-project.org/package=gplots>
140. Löytynoja A. Phylogeny-aware alignment with PRANK. In: Russell DJ, editor. *Multiple Sequence Alignment Methods*. Totowa, NJ: Humana Press; 2014. 155–70.
141. Slowikowski K. ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2'. R package version 0.9.1, 2020. <https://CRAN.R-project.org/package=ggrepel>
142. Wickham H. ggplot2: Elegant Graphics for Data Analysis. 2 ed: Springer-Verlag New York; 2016.
143. Yu G. Scatterpie: Scatter Pie Plot. R package version 0.1.8. 2020. <https://CRAN.R-project.org/package=scatterpie>
144. Wilke C. Cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'. R package version 1.0.0, <https://cran.r-project.org/package=cowplot>. 2019.

145. Dunnington D, Thorne B, Hernangómez D. ggspatial: Spatial Data Framework for ggplot2. R package v. 1.1.1. 2020. <https://cran.r-project.org/package=ggspatial>
146. Pebesma E. Simple features for r: standardized support for spatial vector data. R J. 2018;10(1):439. <https://doi.org/10.32614/rj-2018-009>
147. Bivand R, Rundel C. Rgeos: Interface to Geometry Engine - Open Source ('GEOS'). R package v. 0.5-1, <https://cran.r-project.org/package=rgeos>. 2019.
148. Elff M. Memisc: Management of Survey Data and Presentation of Analysis Results. R package version 0.99.27. 2020. <https://cran.r-project.org/package=memisc>
149. Venables B, Hornik K. Oz: Plot the Australian Coastline and States. 2016. <https://cran.r-project.org/package=oz>
150. Bivand R, Lewin-Koh N. maptools: Tools for handling spatial objects. 2020.
151. South A. rnaturalearth: World Map Data from Natural Earth. R package version 0.1.0. 2017. <https://cran.r-project.org/package=rnaturalearth>.
152. Cooley D. googleway: Accesses Google Maps APIs to Retrieve Data and Plot Maps. 2018. <https://CRAN.R-project.org/package=googleway>
153. Minsavage GV. Plasmid-Mediated Resistance to Streptomycin in *Xanthomonas campestris* pv. *vesicatoria*. Phytopathology. 1990;80(8):719. <https://doi.org/10.1094/phyto-80-719>
154. Staskawicz B, Dahlbeck D, Keen N, Napoli C. Molecular characterization of cloned avirulence genes from race 0 and race 1 of *Pseudomonas syringae* pv. *glycinea*. J Bacteriol. 1987;169(12):5789–94. <https://doi.org/10.1128/jb.169.12.5789-5794.1987> PMID: [2824447](https://pubmed.ncbi.nlm.nih.gov/2824447/)