A Multilocus Sequence Tying Scheme for Rapid Identification of Xanthomonas citri Based on Whole Genome Sequencing Data

Article *in* Phytopathology · April 2024

DOI: 10.1094/PHYTO-12-23-0490-R

6 authors, including:

Efenaide B Okoh
Western Sydney University
3 PUBLICATIONS **0** CITATIONS
SEE PROFILE

Michael John Payne
UNSW Sydney
55 PUBLICATIONS **370** CITATIONS
SEE PROFILE

Markus Riegler
Western Sydney University
330 PUBLICATIONS **8,196** CITATIONS
SEE PROFILE

Daniel Bogema
Elizabeth Macarthur Agricultural Institute
100 PUBLICATIONS **828** CITATIONS
SEE PROFILE

Phytopathology® • 2024 • 114:1480-1489 • https://doi.org/10.1094/PHYTO-12-23-0490-R

Bacteriology

A Multilocus Sequence Typing Scheme for Rapid Identification of *Xanthomonas citri* Based on Whole-Genome Sequencing Data

Efenaide B. Okoh,[1,2] Michael Payne,[3] Ruiting Lan,[3] Markus Riegler,[1] Toni A. Chapman,[2] and Daniel R. Bogema[2,†]

[1] Hawkesbury Institute for the Environment, Western Sydney University, Locked Bag 1797, Penrith, NSW 2751, Australia
[2] Elizabeth Macarthur Agricultural Institute, NSW Department of Primary Industries, Menangle, NSW 2568, Australia
[3] School of Biotechnology and Biomolecular Science, University of New South Wales, Sydney, NSW, Australia
Accepted for publication 24 April 2024.

Abstract

Xanthomonas citri is a plant-pathogenic bacterium associated with a diverse range of host plant species. It has undergone substantial reclassification and currently consists of 14 different subspecies or pathovars that are responsible for a wide range of plant diseases. Whole-genome sequencing (WGS) provides a cutting-edge advantage over other diagnostic techniques in epidemiological and evolutionary studies of *X. citri* because it has a higher discriminatory power and is replicable across laboratories. WGS also allows for the improvement of multilocus sequence typing (MLST) schemes. In this study, we used genome sequences of *Xanthomonas* isolates from the NCBI RefSeq database to develop a seven-gene MLST scheme that yielded 19 sequence types (STs) that correlated with phylogenetic clades of *X. citri* subspecies or pathovars. Using this MLST scheme, we examined 2,911 *Xanthomonas* species assemblies from NCBI GenBank and identified 15 novel STs from 37 isolates that were misclassified in NCBI. In total, we identified 545 *X. citri* assemblies from GenBank with 95% average nucleotide identity to the *X. citri* type strain, and all were classified as one of the 34 STs. All MLST classifications correlated with a phylogenetic position inferred from alignments using 92 conserved genes. We observed several instances where strains from different pathovars formed closely related monophyletic clades and shared the same ST, indicating that further investigation of the validity of these pathovars is required. Our MLST scheme described here is a robust tool for rapid classification of *X. citri* pathovars using WGS and a powerful method for further comprehensive taxonomic revision of *X. citri* pathovars.

Keywords: classification, MLST, scheme, WGS, *Xanthomonas citri*

Xanthomonas is a diverse genus of phytopathogenic bacteria that causes severe diseases in at least 124 monocotyledonous and 268 dicotyledonous plant species worldwide (An et al. 2020; Hayward 1993). The species *X. citri* can infect a wide range of host plant species typically distinguished by host-associated diseases (e.g., *X. citri* subsp. *citri* causes citrus canker). However, host overlaps between pathovars occur (e.g., *X. citri* pv. *anacardii* and *X. citri* pv. *mangiferaeindicae* both infect mango). However, within the *X. citri* subgroups (which are classified as either subspecies or pathovars), there is a restricted host range that members of each subgroup are capable of infecting (Jacques et al. 2016). For clarity, "subspecies or pathovars" are not used for the remainder of this manuscript, and we have referred to all subgroups as pathovars.

Exotic and widespread plant pests and pathogens pose a threat to food security and the long-term viability of agriculture (Pimentel et al. 2001; Vurro et al. 2010). The causal agent of citrus canker is a pathogen that has been known worldwide for the last 20 decades, causing tremendous economic losses amounting to billions of dollars spent in eradication in both the United States and Australia (Graham et al. 2004). However, less is known about the genetic relatedness of other *X. citri* pathovars that could emerge as devasting pathogens in the future.

Classically, characterization of xanthomonads was done using pathogenicity profiling, which usually involves phenotypic identity-based methods (Sintchenko et al. 2007) to confirm and identify if an isolate is pathogenic or nonpathogenic in a particular plant host. Although this is a useful and necessary tool, methods can be complex and labor-intensive, provide limited information on microbial diversity, and can only be performed on a few plant species at a time (Vauterin et al. 1995). Additionally, the diversity of xanthomonads observed with this method led to a "new host-new species" concept (Starr 1981), resulting in extreme nomenclature and classification changes over several decades. Furthermore, epiphytic growth occurs in *X. citri*, leading to potential misclassifications if host species are primarily used as classifiers (Zarei et al. 2018).

More recently, molecular methods such as genomics and phylogenetic analyses have classified xanthomonads at the species level (Parkinson et al. 2009). However, often, *X. citri* subgroups are determined by host and pathogenicity profiling for pathovar classifications, as these often form discrete phylogenetic groups of high sequence similarity (Dubrow et al. 2022). High-throughput (a.k.a. next-generation) sequencing methods have revolutionized the characterisation of bacterial isolates and are likely to replace current typing methodologies due to higher resolution and reduce costs (Sabat et al. 2013). Currently, genome typing schemes developed for outbreak investigations of human-pathogenic bacteria (Payne et al. 2020) can be employed to evaluate epidemiology, evolution (Ragupathy et al. 2023), and genetic relationships of plant pathogenic bacteria.

Multilocus sequence typing (MLST) was first developed for *Neisseria meningitidis* to overcome the poor reproducibility between laboratories using older molecular typing schemes (Maiden et al.

†Corresponding author: D. R. Bogema; daniel.bogema@dpi.nsw.gov.au

Funding: Support was provided by the Grains Research and Development Corporation through funding from the Australian Government Department of Agriculture, Fisheries & Forestry, as part of its Rural R&D for Profit Program and along with Cotton Research and Development Corporation, Hort Innovation Australia, Wine Australia, Sugar Research Australia, and Forest and Wood Products Australia (grant 9177866). E. B. Okoh was supported by a Ph.D. scholarship from Western Sydney University. This work was also funded by the Centre for Invasive Species Solutions through a grant titled "Enhanced preparedness to diagnose 200+ priority pests and diseases of NSW and Australian Agriculture" (A-023).

***e*-Xtra:** Supplementary material is available online.

The author(s) declare no conflict of interest.

1998). Since then, MLST has rapidly become the gold standard for identifying and classifying bacteria in a comprehensible and comparable approach on a global scale. The principle behind this approach is to determine the genetic similarities and relatedness within bacterial species using the DNA sequence information of between four and eight loci for discrimination between strains and identification of clonal lineages (Maiden 2006). For each locus, sequence variants are considered separate alleles, and each is assigned a number. The combination of allele numbers of all loci results in an allelic profile for each strain (also known as sequence type [ST]). Housekeeping genes are widely used in MLST schemes because they encode central metabolic functions and are conserved within the same species. Mutations within these genes are largely synonymous and selectively neutral, therefore they are more likely to properly reflect the phylogeny of strains (Gevers et al. 2005). Moreover, the use of multiple loci provides a buffer against the distorting effect of recombination (Gevers et al. 2005).

Originally, conventional MLST relied on PCR amplification and Sanger sequencing of several housekeeping genes. However, whole-genome sequencing (WGS) data are now often used as a simple and fast approach for in silico MLST to establish strain identity, as allele profiles can be extracted from the genome sequences (Bezdicek et al. 2023). Overall, MLST enables easy comparison of bacterial strains and relatedness of an isolate that cannot be inferred from phenotypic or other molecular typing methods (Bezdicek et al. 2023). Although a core genome MLST scheme for *X. citri* has been completed (Ragupathy et al. 2023), conventional MLST is yet to be applied to *X. citri*.

Here, we developed a seven-gene MLST scheme to characterize *X. citri* isolates for rapid identification and classification of *X. citri* pathovars based on WGS or gene sequencing data. Our study also provides valuable genomic information for the development of molecular techniques aimed at identifying specific pathogens and enhancing detection methods that may be effective for disease management in agricultural systems.

Materials and Methods

Sequence data, phylogenetic identification, and core gene identification

Initially, to demonstrate the relationship between *X. citri* pathovars and sequence similarity, we generated a scheme construction dataset of 163 publicly available *X. citri* genomes downloaded from the NCBI RefSeq database (Table 1). We examined the phylogenetic position of the pathovars using 92 conserved loci extracted from assemblies with UBCG v3. This program works by searching genomes for a set of 92 core genes that are conserved across all bacteria. These genes are then aligned, and the alignments are concatenated into a supermatrix alignment for phylogenetic analysis (Na et al. 2018). A maximum likelihood tree was constructed from the UBCG-generated supermatrix alignment using RAxML v8.2.12 and the GTRGAMMA model and -b 123 (Stamatakis 2014). Phylogenetic trees were visualized using Interactive Tree of Life (iTOL) v5 (Letunic and Bork 2021).

To identify core genes, genes were first annotated with Prokka v1.14.6 (Seemann 2014) to determine those present in all scheme construction dataset assemblies. Then, annotated assemblies were used to determine the core loci of *X. citri* using Panaroo v1.3.3 (Tonkin-Hill et al. 2020) using default parameters, including the *–merge_paralogs*, *–clean-mode strict*, and *–remove-invalid-genes* options.

We identified true core genes and removed paralogues by using a custom Python script (Script 1, https://github.com/ngraned/Scripts-Seven-Gene-MLST-*Xanthomonas-citri*) to find the common genes in the entire dataset. Then, the panaroo-filter-pa script from Panaroo was used to filter out pseudogenes. Finally, genes with premature stop codons (Script 2, https://github.com/ngraned/Scripts-Seven-Gene-MLST-*Xanthomonas-citri*) and paralogues were removed using a custom Python script (https://github.com/LanLab/NestedProkkaFix). Alleles for all remaining core genes were classified and assigned allele numbers for each assembly using an established genome typing pipeline (Payne et al. 2020). Grapetree v1.5.0 (Zhou et al. 2018) was used to visualize each gene after the alleles were assigned.

MLST gene evaluation

The chromosomal positions of MLST loci were assessed using BRIG v0.95 (Alikhan et al. 2011). The ratio of nonsynonymous to synonymous substitutions (dN/dS) was calculated to determine if genes were under positive selection pressure using the method established by Nei and Gojobori (1986). To complete the dN/dS analysis, multiple sequence alignments of 50 representative genomes (a subset of the 163 genomes scheme construction dataset) were used to reduce the computational power required for analysis. Alignments of each locus were generated using MAFFT v7.520. Codon alignments were built using Pal2nal v14, and dN, dS, and dN/dS ratios were calculated using the Synonymous Nonsynonymous Analysis Program (SNAP) v2.1.1 (Korber 2000).

We used different databases and tools to assess the chosen gene functional categories for the likelihood of positive selection pressure (Supplementary Table S2). Genes core to the *Xanthomonas* genus were identified using 25 *Xanthomonas* spp. from phylogenetic clade 2 identified with a maximum-likelihood tree constructed using FastTree v2.1.11 (Price et al. 2010) (Supplementary Fig. S1). Genes core to the *Xanthomonas* genus were then identified using Roary v3.13.0 (Page et al. 2015) with default parameters and *-i* (minimum percentage identity for blastp) set to 70.

Validation of the MLST scheme

To assess the specificity of the MLST scheme, a specificity dataset consisting of 2,911 publicly available *Xanthomonas* spp. genome assemblies (Table 1) was examined by MLST using profiles generated from the 163-assembly scheme construction dataset. MLST profiles were generated from the specificity dataset using mlst v2.19.0 (https://github.com/tseemann/mlst). *X. citri* assemblies were identified from the specificity dataset using organism taxonomy linked to each GenBank assembly and average nucleotide identity (ANI). Assemblies classified as *X. citri* by NCBI and with a 95% ANI value when queried to the *X. citri* type strain A306 (GCA_000816885.1) were added to the *X. citri* positive dataset (Table 1). ANI values were calculated using fastANI v1.33 (Jain et al. 2018). Likely pathovars of each *X. citri* positive dataset assembly were determined by phylogenetics using UBCG v3 (Na et al. 2018) and RAxML v8.2.12 as described above. The GenBank accession numbers of all genomes used in this study are listed in Supplementary Table S1.

DNA sequences determined in this study

The DNA sequences of the alleles analyzed in this study were obtained from the genome sequences available with the provided accession numbers at NCBI. The allele sequences have been

TABLE 1. Sequence datasets used in this study

Dataset name	Assemblies	Source	Date collected	Explanation
Scheme construction	163	RefSeq	September 2021	Define pathovars to build multilocus sequence typing scheme
Specificity	2,911	GenBank	May 2023	*Xanthomonas* spp. assemblies used to test the designed scheme
Positive *X. citri*	545	GenBank	May 2023	Assemblies from specificity dataset confirmed to be *X. citri* by average nucleotide identity

provided as a supplement to this manuscript (Supplementary Data S1).

Results

Phylogenetics and pathovars are related in *X. citri*

Previous studies have shown that *Xanthomonas* spp. pathovars can be identified by phylogenetic analysis (Parkinson et al. 2009). We phylogenetically examined the genome assemblies of the scheme construction dataset and compared the results with their pathovar classification sourced from NCBI (Fig. 1) to demonstrate that this is also the case for *X. citri*. Most isolates in pathovar groups, such as *citri*, *fuscans*, *malvacearum*, and *mangiferaindicae* genomes, were correctly assigned to the pathovar level from NCBI and match the phylogenetic pathovar assignment. Only a few isolates were assigned to the species level of *X. citri*, and these isolates (red font in Fig. 1) can be easily identified to the pathovar level based on their clustering with the clades of known pathovars.

MLST gene selection

We chose seven genes based on the association between allele profile and phylogeny-pathovar classification and positions within the chromosome of our chosen reference genome *X. citri* pv. *citri* MN12 (Fig. 2). These were *nusG*, *rpmA*, *arfB*, *rplQ*, *queD*, *fdx*, and *exbD* (Table 2). All loci were confirmed to be single-copy genes, and their complete sequences were present in more than 95% of all *X. citri* GenBank genomes.

We wanted to determine if these genes were good candidates for an MLST scheme. Therefore, we examined these loci with several bioinformatic methods to identify the likelihood that these genes were undergoing positive selection. The calculated ratio of synonymous and nonsynonymous substitutions (dN/dS ratio) of the seven chromosomal genes chosen indicated that these genes were not under positive selection pressure (Supplementary Table S3). In addition to the dN/dS ratio, we tested each gene using a set of characteristics, such as the presence of transmembrane helices, Clusters of Orthologous Genes category, phage genes, homopolymers, and tandem repeats to further determine the likelihood of positive selection and others listed in Supplementary Table S2. We found that these genes were unlikely to be undergoing positive selection. However, one gene, *exbD*, did have some aspects not related to housekeeping genes (transmembrane helix and Clusters of Orthologous Genes category U—intracellular trafficking, secretion, and vesicular transport). We examined other schemes in pubMLST and found that the presence of transmembrane helices was common. This gene was kept because it was highly associated with pathovar classification and there was not a more suitable alternative to the wider MLST scheme.

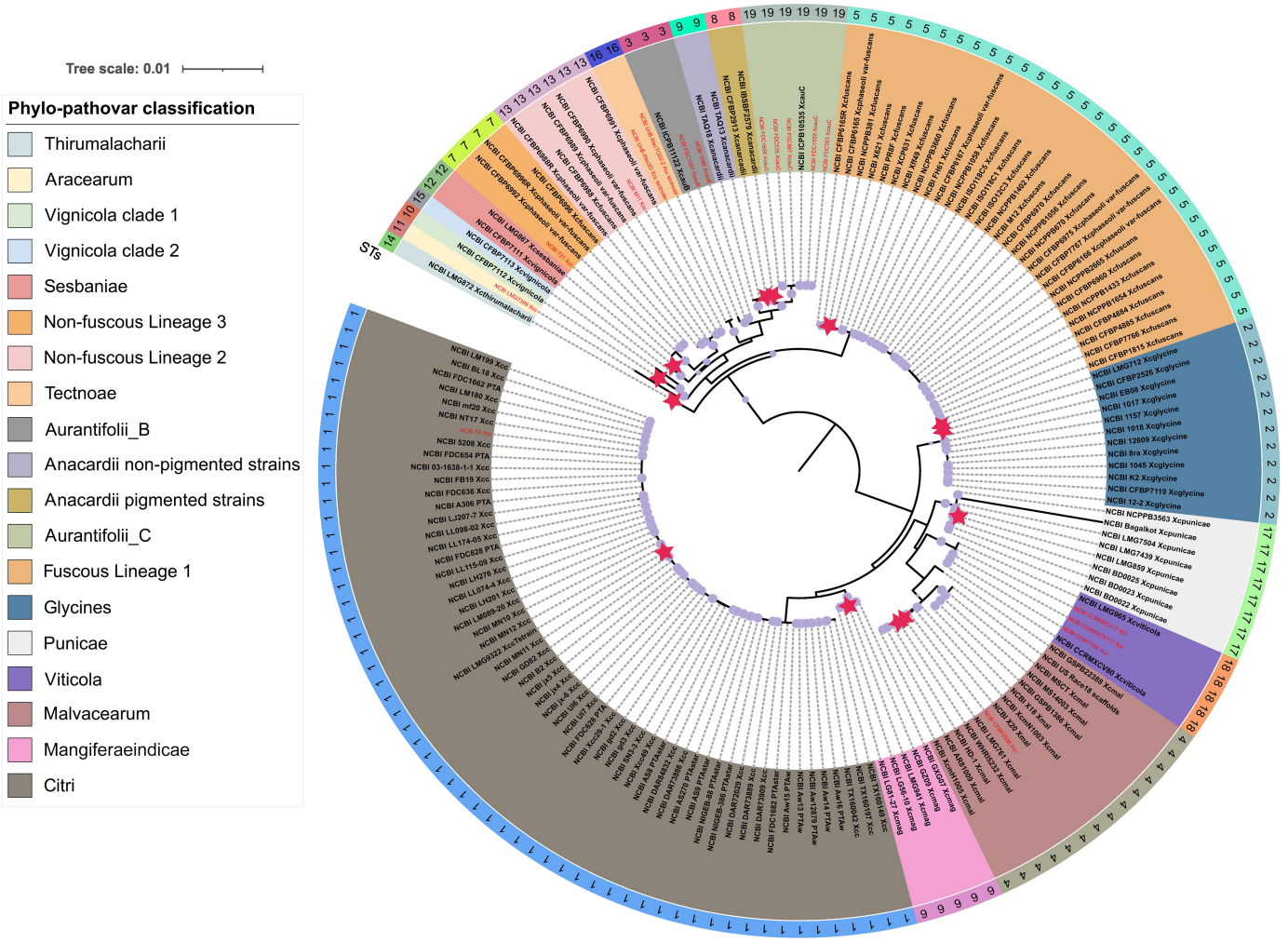

Fig. 1. Maximum-likelihood phylogeny was used to build the multilocus sequence typing scheme with 163 NCBI genomes using RAxML (Stamatakis 2014). Nineteen sequence type (ST) classifications were matched with each pathovar group identified and represented in the phylo-pathovar color legend. The bold black font indicates assemblies assigned to the pathovar level, and the red font indicates the genomes identified to the *Xanthomonas citri* level by NCBI. Genomes marked with red stars are the pathovar type strains, and purple circles mark genomes with bootstrap supported values above 70%.

Specificity validation using 2,911 *Xanthomonas* spp. genomes

From the scheme construction dataset, we identified 19 STs, each exclusively associated with a pathovar of *X. citri*. Some pathovars (*fuscans*, *aurantifolii*, *anacardii*, and *vignicola*) had more than one ST, but these were associated with defined phylogenetic groups within these clades (Fig. 1). Each of the initial 19 STs identified had a minimum of two allele differences from all other profiles in the scheme, meaning that if any one gene is partial or missing, then the ST uncertainty is likely to be rare and pathovar can still be assigned (Table 3).

We wanted to test the MLST scheme against genome sequences of varying quality and to examine if the STs were unique to *X. citri*. Therefore, we used the constructed scheme to profile the 2,911 *Xanthomonas* GenBank assemblies of the specificity dataset (Table 1). We identified 545 assemblies as positive *X. citri* by 95% ANI queried using *X. citri* pv. *citri* A306. The MLST scheme returned complete MLST profiles for 505 assemblies; 19 assemblies were assigned 15 novel STs (Table 3). All 524 assemblies matched the phylogenetic positions expected by their called ST, and all assemblies of novel STs formed monophyletic clades or single tips in phylogenetic trees (Fig. 3; Supplementary Fig. S2).

To assess our MLST scheme performance further, we examined the 19 assemblies from the specificity dataset with novel MLST profiles. These assemblies represent different *Xanthomonas* species listed among the NCBI organism names (Table 3). The majority of these pathovars have also been proposed to be reclassified as *X. citri* pathovars by Bansal et al. (2022), except for strains of pv. *olitorii*, pv. *eucalyptorum*, and pv. *phaseoli*.

Amongst the 15 novel STs, two assemblies classified in NCBI as *X. campestris* pv. *vitiscarnosae* and *X. axonopodis* pv. *martynicola* shared the same ST and clustered together phylogenetically (Fig. 3). *X. campestris* pv. *azadirachtae* LMG 543 shared six out of

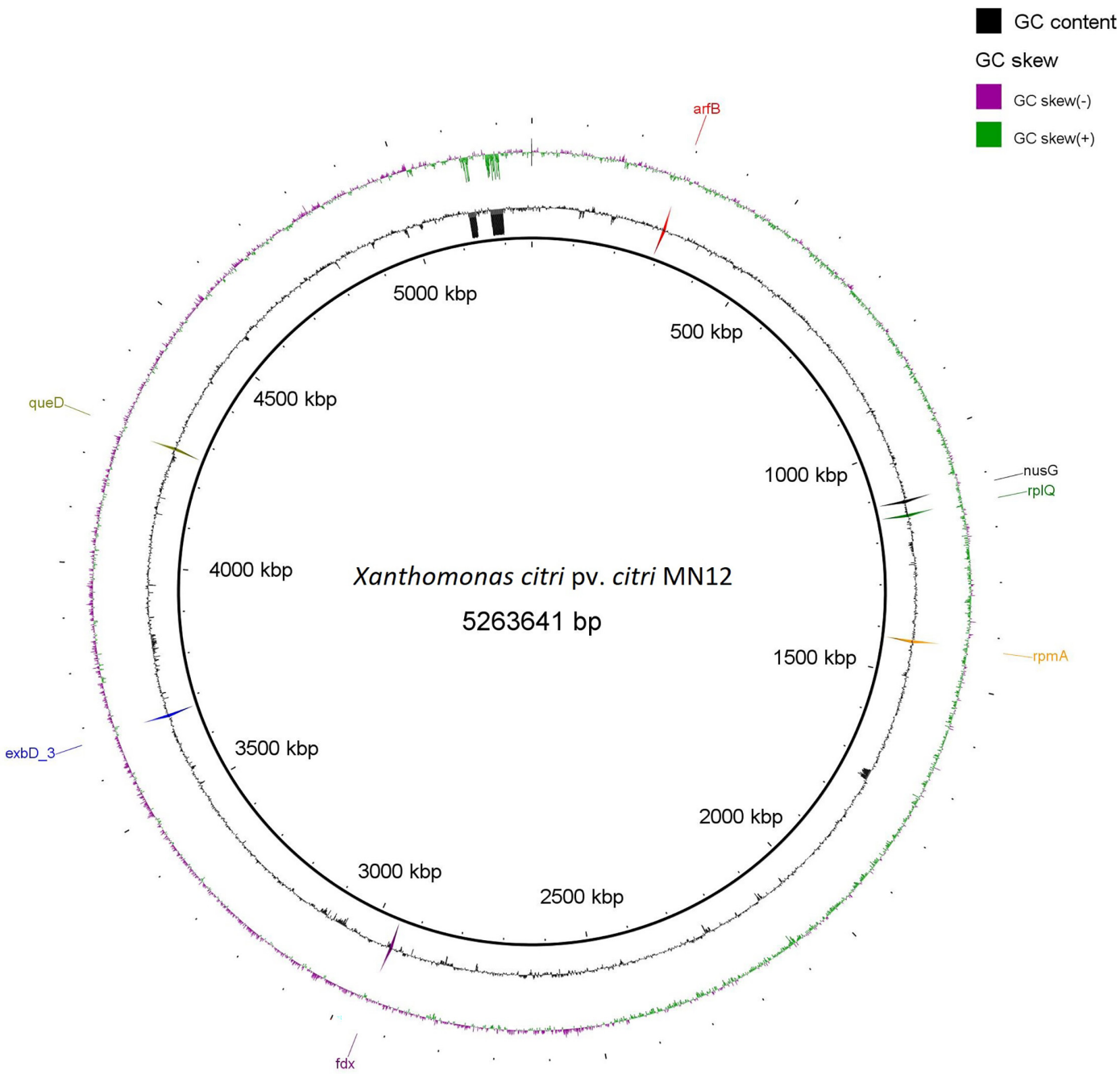

Fig. 2. Chromosomal positions of the seven genes chosen for multilocus sequence typing using *Xanthomonas citri* pv. *citri* MN12 as the reference genome.

seven alleles with *X. campestris* pv. *bilvae* NCPPB 3213. Phylogenetic analysis placed both these assemblies in a single monophyletic clade (Table 3; Fig. 3).

Sixteen assemblies not classified by NCBI as *X. citri* within the 505 assemblies were assigned STs from the 19 defined STs that matched the *X. citri*-associated pathovars (Table 4). Two assemblies (GCA_020879475.1 and GCA_020879395.1), classified by NCBI as *X. citri* pv. *fuscans*, gave full MLST profiles that matched *X. citri* pathovars *vignicola* and *aurantifolii*, respectively (Table 4). For all 18 assemblies, the phylogenetic tree position matched the classified ST and not the given organism's name in NCBI (Table 4; Fig. 3).

In addition, of the total 545 *X. citri* assemblies, 21 had partial allele profiles with up to two loci partially sequenced or missing. These were still typed by MLST, as the scheme is robust enough to accept one to two partial or missed genes (Supplementary Table S4). All 21 assemblies with partial allele profiles from our MLST scheme matched the predicted ST and associated phylogenetic position (Fig. 3). Two assemblies with NCBI organism name *X. citri* pv. *mangiferaeindicae* were not classified by MLST and

TABLE 2. Functional characteristics of the seven genes from reference genome *Xanthomonas citri* pv. *citri* MN12 used to construct the multilocus sequence typing scheme

Gene name	Gene position	Gene size (bp)	Function of gene product	Number of alleles	Number of alleles with ambiguous bases	Number of alleles with incomplete sequence
nusG	1117908−1118465	558	Transcription termination/antitermination protein	22	0	0
rpmA	1425831−1426091	261	50S ribosomal protein L27	18	0	0
arfB	293853−294284	432	Peptidyl-tRNA hydrolase	22	1	0
rplQ	1148966−1149349	384	50S ribosomal protein L17	9	0	0
queD	4265629−4265991	363	6-Carboxy-5678-tetrahydropterin synthase	14	0	0
fdx	2947012−2947296	285	Ferredoxin	23	0	0
*exbD*_3	3671473−3671922	450	Biopolymer transport protein	19	0	0

TABLE 3. Properties of 505 isolates from the dataset of 545 with 19 defined sequence types (STs) and complete allele profiles of the multilocus sequence typing scheme and novel allele profiles of 19 misclassified genomes from the NCBI dataset of 545 identified as *Xanthomonas citri*

Pathovar	Number of isolates	ST	*nusG*	*rpmA*	*arfB*	*rplQ*	*queD*	*fdx*	*exbD*	Minimum number of allele distance to another ST
X. citri pv.										
citri	355	1	1	1	1	1	1	1	1	4
glycines	19	2	2	2	2	2	1	2	2	5
*aurantifolii*_B	3	3	3	3	3	3	2	3	3	3
malvacearum	16	4	4	4	4	4	1	4	4	4
fuscans lineage 1 (L1)	35	5	5	5	3	5	3	3	5	4
mangiferaindicae	6	6	6	2	5	1	1	1	6	4
Non-fuscous L3	7	7	7	6	3	6	4	5	5	3
anacardii pigmented	9	8	8	3	9	3	5	6	3	3
anacardii non-pigmented	2	9	8	6	3	3	5	6	5	3
vignicola-clade 1	3	10	8	6	3	6	7	11	5	3
aracearum[a]	5	11	8	6	3	7	10	13	11	4
sesbaniae	4	12	8	6	8	1	3	10	5	2
Non-fuscous L2	11	13	8	6	8	5	3	9	5	2
thirumalacharii	1	14	8	6	10	5	12	6	3	4
vignicola-clade 2	1	15	8	9	3	5	8	12	9	5
tectnoae[b]	2	16	8	10	8	5	11	14	3	4
punicae	14	17	9	7	6	2	1	7	7	5
viticola	5	18	10	8	7	1	6	8	8	3
*aurantifolii*_C	7	19	11	3	3	3	9	3	10	3
Novel profiles										
X. campestris pv. *olitorii* NCPPB 464	1	20	12	7	11	5	1	15	12	5
X. axonopodis pv. *eucalyptorum* LPF 602	1	21	8	3	12	3	3	3	3	3
X. campestris pv. *vitiswoodrowii* LMG 954	1	22	13	11	13	8	6	16	13	5
X. cissicola LMG 21719 and CCUG 18839	2	23	14	12	14	9	6	8	8	4
X. campestris pv. *bilvae* NCPPB 3213	1	24	15	13	15	1	1	1	14	1
X. phaseoli pv. *phaseoli* NCPPB 556, NCPPB 2064, and NCPPB 557 (proposed name: pv. *lablab* clade 2)	3	25	16	14	16	1	13	17	15	4
X. axonopodis pv. *khayae* LMG 753	1	26	17	15	3	1	6	18	8	4
X. campestris pv. *vitistrifoliae* LMG 940	1	27	18	11	7	1	6	19	8	3
X. campestris pv. *vitiscarnosae* LMG 939 and *X. axonopodis* pv. *martyniicola* LMG 9049	2	28	10	16	17	1	6	8	16	3
X. axonopodis pv. *bauhiniae* LMG 548	1	29	19	17	18	1	6	20	17	5
X. axonopodis pv. *cajani* LMG 558	1	30	20	18	19	1	1	21	18	4
X. axonopodis pv. *melhusii* LMG 9050	1	31	21	4	4	5	1	8	8	4
X. axonopodis pv. *clitoriae* LMG 9045	1	32	22	1	20	1	1	22	18	4
X. campestris pv. *azadirachtae* LMG 543	1	33	15	13	21	1	1	1	14	1
X. phaseoli pv. *phaseoli* NCPPB 1713 (proposed name: pv. *lablab* clade 1)	1	34	16	14	22	1	14	23	19	4

[a] The isolates are classified as pv. *aracearum* in this study as proposed by Constantin et al. (2016).
[b] The isolates are classified as pv. *tectnoae* in this study as proposed by Monteiro et al. (2023).

were confirmed not to be *X. citri* by ANI (Supplementary Table S4). We further examined NCBI records for these assemblies and found that they do not show greater than 95% ANI to any known bacterial species and are unlikely to be sourced from *Xanthomonas* species.

Discussion

Improved understanding of the relationships among pathogen variants and strains through molecular typing is one of the fundamental components of modern epidemiology (Ranjbar et al. 2014; Simar et al. 2021). Reliable classification and precise strain characterization are needed to identify pathogens rapidly and accurately in the case of an epidemic outbreak to find their genetic relationships, distribution, epidemiology, and evolutionary changes (Ragupathy et al. 2023). The integration of recent molecular approaches such as WGS as a faster and cheaper option has reshaped bacterial strain typing and brought a revolutionary shift in the field of epidemiology. Here, we developed an in silico MLST scheme based on WGS data from all available *X. citri* genomes worldwide to infer the phylogenetic and population dynamics of *X. citri* pathovars.

X. citri comprises a genetically proximate cluster of strains that diversify into different clonal lineages or sublineages (Midha and Patil 2014). Based on the data generated using the MLST scheme presented here, we suggest that *X. citri* pathovars constitute clearly delineated genomic species (Parkinson et al. 2007; Rademaker et al. 2005; Vauterin et al. 1995; Young et al. 2008). All alleles at each gene were specific to *X. citri* when compared with other species of the genus *Xanthomonas* (Supplementary Table S1). Additionally, we demonstrate that (i) our MLST scheme is capable of clearly defined typing of *X. citri* pathovars with STs correctly assigned to

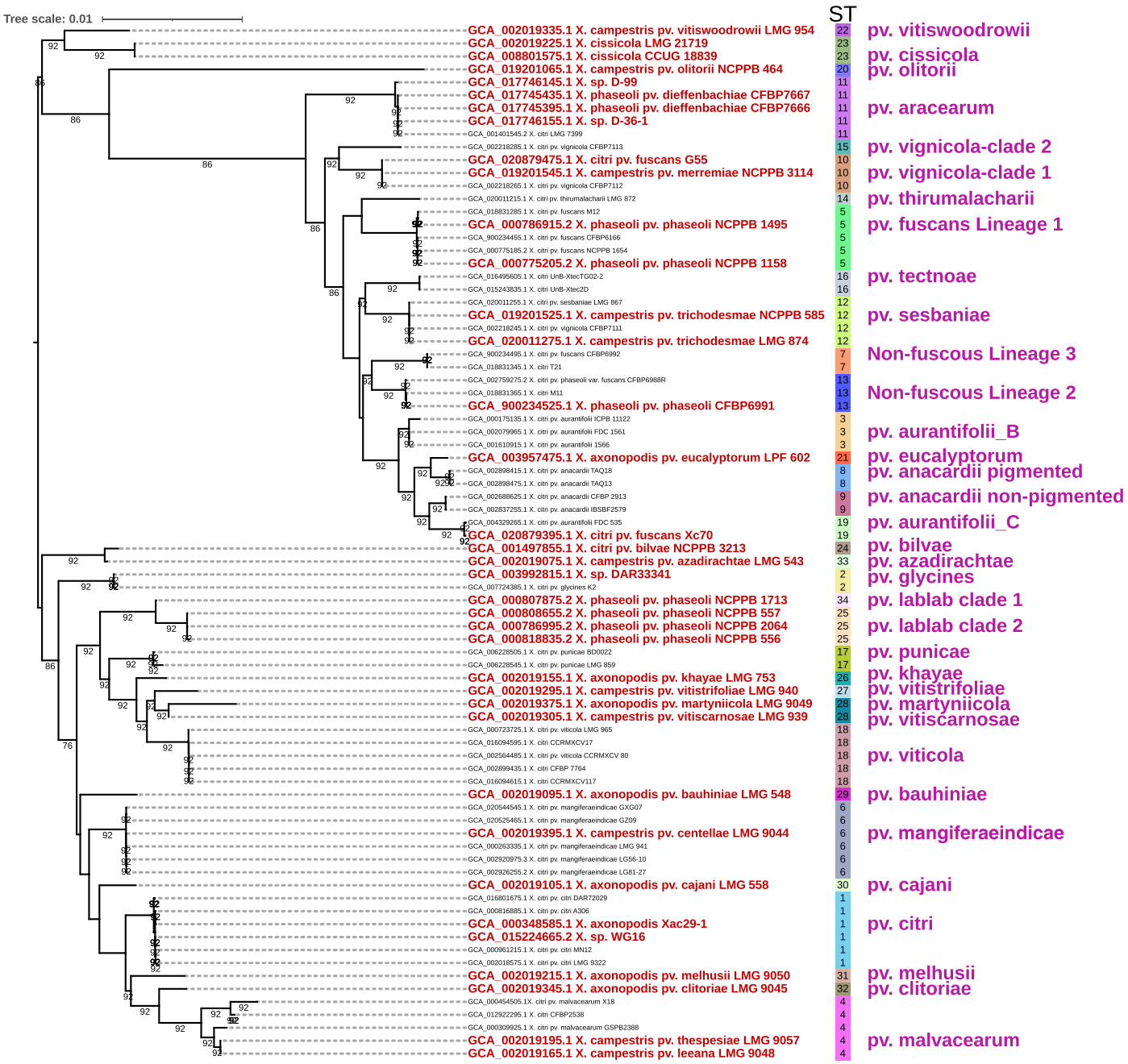

Fig. 3. Multilocus sequence typing profile of 80 selected representative genomes of *Xanthomonas citri*, together with a maximum-likelihood phylogeny using RAxML (Stamatakis 2014). Bootstrap support values are above 70%. The bold red font represents the 37 genomes with incorrect NCBI classification, and sequence type represents the association with pathovar call.

pathovars, (ii) the chosen loci are polymorphic and not subjected to positive selection, and (iii) the chosen loci are typeable, stable, and present in all *X. citri* strains.

In our study, STs were identified within *X. citri* pathovars, indicating a high degree of genetic diversity. Previous genome sequencing studies on the evolution and phylogeny of *X. citri* (Gordon et al. 2015; Timilsina et al. 2020) have shown the diversity of *X. citri* pv. *citri* at high resolution. Overall, these studies suggest that *X. citri* has undergone significant evolutionary changes. Previous studies have observed that strains of *X. citri* pv. *anacardii* isolated from both cashew and mango hosts are separated into pigmented and non-pigmented strains, which are phylogenetically distinct (da Gama et al. 2018; Gama et al. 2011; Lucena et al. 2023; Silva Junior et al. 2018; Zombre et al. 2016). In this study, MLST differentiated pigmented and non-pigmented strains into two STs based on their allele sequence differences and phylogenetic position. Hence, the selected MLST genes with the complete allele sequences obtained from WGS have more discriminatory power to observe genetic relatedness among isolates of the same pathovar from different host plants. Although two STs were identified with *X. citri* pv. *vignicola*, each strain was genetically unique. We found no clear *X. citri* pv. *vignicola* cluster, although the two strains were closely related to each other, as supported by a high bootstrap value.

Interestingly, strain CFBP7111, classified as *X. citri* pv. *vignicola*, isolated from cowpea (Ruh et al. 2021), phylogenetically clustered most closely to *X. citri* pv. *sesbaniae* type strain LMG 867 isolated from *Sesbania sesban* (Egyptian riverhemp). Both were collected from different continents, America (United States) and Asia (India), and isolated more than a decade apart (1942 and 1958). This unexpected genetic relatedness with high bootstrap value prompts a question about the primary factor behind this divergence as to why this strain clustered within the pathovar *sesbaniae*. This could represent a misclassification. However, little is known about these pathovars, and the *X. citri* pv. *vignicola* pathovar type strain (LMG 8752 same as CFBP7112) has a different phylogenetic position and ST. Further study of legume or pea family xanthomonads would clarify the nomenclature of these isolates.

Similarly, strains NCPPB 585 and LMG 874 classified previously as *X. campestris* pv. *trichodesmae* isolated from camel bush also grouped with *X. citri* pv. *sesbaniae* in both MLST and phylogenetics. Bansal et al. (2022) proposed that *X. campestris* pv. *trichodesmae* be reclassified into *X. citri*. This proposal is supported by the genetic relatedness found between these strains and strains of *X. citri* pv. *sesbaniae*. Reclassification of *X. campestris* pathovars often occurs when molecular analyses (such as MLST and multilocus sequence analysis [MLSA]) show limited genetic differences between pathotype strains (Fargier et al. 2011). This finding shows the need for additional analyses of *X. citri* in legumes, including the confirmation of Koch's postulates for this pathovar.

MLST grouped both *X. campestris* pv. *merremiae* NCPPB 3114, isolated from *Merremia gangetica* in India in 1976, and the misclassified *X. citri* pv. *fuscans* G55, isolated from cowpea, with *X. citri* pv. *vignicola* type strain CFBP7112 (clade 1, cowpea host; Ruh et al. 2021). *Merremia gangetica* (Ganges morning glory) could be a new host for this pathovar. Bansal et al. (2022) proposed the reclassification of *X. campestris* pv. *merremiae* into *X. citri*. We support this proposal based on the phylogenetic position and allelic profile in this study between the strain to pv. *vignicola* CFBP7112. Even though the strains are closely related, due to the limited number of isolates, it is difficult to draw strong conclusions. This finding shows the need for further additional sequencing and testing of Koch's postulates for this pathovar.

Numerous other assemblies of *Xanthomonas* strains given different pathovar names based on their host of isolation had the same MLST profile and therefore high genetic relatedness. Strains received as *X. campestris* pv. *leena* LMG 9048 isolated from *Leea asiatica* (Bandicoot Berry; family *Vitaceae*; isolated 1967) and *X. campestris* pv. *thespesiae* isolated from *Thespesia populnea* (Portia tree; family *Malvaceae*; isolated 1978) were proposed to be reclassified into *X. citri* based on phylogenomic analysis (Bansal et al. 2022). Our findings support this proposal due to the high genetic similarity and relatedness to pv. *malvacearum* in both their MLST profiles and phylogenetic position. These two strains could represent a new group of host plants for pv. *malvacearum*. In a similar context, strains LMG 9049 and LMG 939, isolated from *Martynia annua* (small fruit devil's claw) and *Cayratia trifolia* (bush grape) in 1958 and 1962, respectively, had the same MLST profile and are genetically very similar based on their phylogenetic position. Our finding supports the proposal by Bansal et al. (2022) to reclassify these strains into *X. citri*.

In addition, there may be more pathovars of *X. campestris* that are still misclassified at the species level and will require further revision, such as *X. campestris* pv. *olitorii* NCPPB 464 isolated from *Corchorus olitorius* (jute mallow; family *Malvaceae*) in Sudan, which groups as an *X. citri* pathovar according to ANI, MLST, and phylogenetic analyses. Despite it belonging to the same group

TABLE 4. Sixteen genomes within the 505 assemblies not classified as *Xanthomonas citri* and two genomes within the 505 assemblies classified as an incorrect pathovar of *X. citri*

Accession number	Intraspecific name (strain)	Declared organism name in NCBI	Best matching strain/species name	Average nucleotide identity (type strain A306)	Sequence type and pathovar clustering
GCA_002019165.1	LMG 9048	*X. campestris* pv. *leeana*	*X. citri*	98.64	4-*malvacearum*
GCA_002019195.1	LMG 9057	*X. campestris* pv. *thespesiae*	*X. citri*	98.62	4-*malvacearum*
GCA_002019395.1	LMG 9044	*X. campestris* pv. *centellae*	*X. citri*	98.70	6-*mangiferaindicae*
GCA_019201525.1	NCPPB 585	*X. campestris* pv. *trichodesmae*	*X. citri*	96.12	12-*sesbaniae*
GCA_019201545.1	NCPPB 3114	*X. campestris* pv. *merremiae*	*X. citri*	96.16	10-*vignicola*
GCA_020011275.1	LMG 874	*X. campestris* pv. *trichodesmae*	*X. citri*	96.13	12-*sesbaniae*
GCA_000348585.1	Xac29-1	*X. axonopodis* Xac29-1	*X. citri*	99.95	1-*citri*
GCA_000775205.2	NCPPB 1158	*X. phaseoli* pv. *phaseoli*	*X. citri*	96.17	5-*fuscans*
GCA_000786915.2	NCPPB 1495	*X. phaseoli* pv. *phaseoli*	*X. citri*	96.11	5-*fuscans*
GCA_017745395.1	CFBP7666	*X. phaseoli* pv. *dieffenbachiae*	*X. citri*	95.85	11-*aracearum*
GCA_017745435.1	CFBP7667	*X. phaseoli* pv. *dieffenbachiae*	*X. citri*	95.94	11-*aracearum*
GCA_900234525.1	CFBP6991	*X. phaseoli* pv. *phaseoli*	*X. citri*	96.25	13-non-fuscous L2
GCA_003992815.1	DAR33341	*Xanthomonas* sp. DAR33341	*X. citri*	98.76	2-*glycines*
GCA_015224665.2	WG16	*Xanthomonas* sp. WG16	*X. citri*	99.97	1-*citri*
GCA_017746155.1	D-36-1	*Xanthomonas* sp. D-36-1	*X. citri*	95.86	11-*aracerum*
GCA_017746145.1	D-99	*Xanthomonas* sp. D-99	*X. citri*	95.99	11-*aracerum*
Incorrectly classified as *X. citri*					
GCA_020879475.1	G55	*X. citri* pv. *fuscans*	*X. citri* pv. *vignicola*	96.22	10-*vignicola*
GCA_020879395.1	Xc70	*X. citri* pv. *fuscans*	*X. citri* pv. *anacardii* CFBP2913	96.35	19-*aurantifolii*_C

as pv. *malvacearum* associated with *Malvaceae*, this strain is distinct in ST and phylogenetic position, necessitating the need to test the Koch's postulates for this pathovar to help resolve its host plant pathogenicity. A cross-pathogenic test and other genetic analyses may need to be performed to identify if any differences in the strains exist to determine if such strains are in fact pathogenic variants or if they are the same organisms capable of causing disease in multiple hosts.

Four misclassified strains named as *X. phaseoli* pv. *phaseoli*, all isolated from *Lablab purpureus* (lablab beans, family *Fabaceae*), had different MLST profiles and phylogenetic positions than the other strains associated with *Phaseolus* (*Fabaceae* host family), pv. *fuscans*, non-fuscous lineage 2 (NF2) and NF3. We suggest that the strains associated with *Fabaceae* are a genetically diverse group. Three strains, NCPPB 557 and NCPPB 556 from 1957 and NCPPB 2064 from 1965 isolated from Sudan, had the same ST, and this matched the phylogenetics position. Strain NCPPB 1713, isolated from Zimbabwe in 1962, had a different ST, and this matched the phylogenetic position. It is confirmed that strains collected from lablab beans are a different genetic lineage from the ones infecting the *Fabaceae* family phylogenetically closely associated with strains of pv. *glycines*. This agrees with previous studies by Aritua et al. (2015) that there is an intra-lineage variation within the lablab-associated strains. Further study of the *Fabaceae* xanthomonads would clarify the nomenclature of these isolates.

Strain LPF 602 *X. axonopodis* pv. *eucalyptorum*, isolated from *Eucalyptus grandis* × *Eucalyptus urophylla* (eucalypt hybrid urograndis) in Brazil in 2000, has a novel ST, and this prediction matched its phylogenetic position. This strain is confirmed by 95% ANI in our study to be *X. citri* and is phylogenetically closely associated between pvs. *aurantifolii*_C and *anacardii* pigmented strains. This proposal is contrary to the study by Ferraz et al. (2018). They proposed that the strain belongs to *X. axonopodis* pv. *eucalyptorum* pv. nov. based on a polyphasic approach that included MLSA and biochemical and pathogenic characteristics. However, it remains inconclusive as to where this strain belongs. This is because we have used ANI, a more sensitive analytical tool for genomic comparisons.

Evidence from previous epidemiological and molecular studies, such as DNA-DNA hybridization, amplified fragment length polymorphism, rep-PCR (Rademaker et al. 2005), and MLSA (Young et al. 2008), support the coherent classification of *Xanthomonas* by Vauterin et al. (1995) concluding that there are several species with limited host ranges and a genetically diverse group affecting numerous host species. Within the *X. citri* group, it remains unclear whether some strains (with diverse host associations) should be their own pathovar or be among the same pathovar group based on genetic relatedness and phylogenetic analyses. Thereby, distinguishing species from closely related genomes may yield less robust descriptions. A preferable approach may be to recognize these groups of strains as pathovars within their phylogenetic clusters until a comprehensive evaluation of *X. citri* is conducted based on further transmission studies. Arbitrarily creating species from a heterogeneous group is likely to necessitate further revisions; this is because it can lead to confusing nomenclature (Young et al. 2008) and result in a reversion back to the "new host-new species" concept. The use of this concept previously led to a swift proliferation in the quantity of phytopathogen species, leading to the formation of complex phytopathogen genera containing hundreds of species.

Whole-genome sequences would be the most informative in determining the overall taxonomic relatedness as predicted by Hauben et al. (1997). Similarly, Zeigler (2003) observed that overall genome relatedness with accurate precision can be predicted using whole-genome sequences. MLSA and MLST studies have used multiple loci to type the genus *Xanthomonas* or delineate new species into *X. citri* pathovars (Almeida et al. 2010; Bansal et al. 2022; Constantin et al. 2016; Fargier et al. 2011; Young et al. 2008) as an alternative approach to DNA-DNA hybridization. These schemes combine different categories of loci that are likely to evolve at different rates. However, using loci that evolve at different rates may pose problems in inferring phylogenetic trees (Cunningham 1997; Matsen et al. 2008).

Young et al. (2008) suggested that the choice of housekeeping genes will need to be refined further for typing purposes, when more complete chromosomal sequences are available. Our MLST scheme developed here has robustly refined the choice of genes and yielded the large diversity of data validated here. The genes are likely to be stable in new types and can be used confidently to infer the rapid classification of *X. citri* pathovars. We have developed a WGS MLST scheme that provides the ability to target conserved loci across the entire core genome to find the most appropriate locus combinations. Our approach maximizes the discriminatory power and brings the resulting ST clustering as close as possible to the WGS analysis results and to the real strain divergence. This offers several advantages, such as molecular comparisons of recent and historical strains with greater precision and accuracy.

Our newly proposed scheme uses the power of WGS to select genes for MLST but also preserves MLST for routine microbial typing, which is still considered the gold standard and offers the possibility of interlaboratory comparisons. Despite the increasing accessibility of WGS, many routine laboratories still lack the necessary resources to fully utilize its potential. Nanopore sequencing has the potential to fill this gap, as the substantially lower equipment costs enable laboratories with fewer resources to perform this type of research. One advantage of in silico MLST such as this *X. citri* MLST scheme is that all genes chosen lack long (>8 bp) homopolymer regions. Insertion-deletion errors near homopolymers are the most common error type in Nanopore sequencing, meaning that our scheme could be utilized widely with low-cost sequencing platforms. These chosen MLST genes can be used to develop a multiplex rapid assay to maintain an integral part of epidemiological investigations and make use of new available sequencing technologies as a fast and cheaper option.

To conclude, our study clarifies genetic relationships and rapidly identifies strains of *X. citri* causing diseases in a wide host plant range and shows that pathovars infecting the same host plant family may be genetically different. We recommend that a revision of *X. citri* pathovars be completed, as our study highlights inconsistences in pathovar classification. The 37 misclassified genomes identified here may have consequences not just for pvs. *citri* and *fuscans* listed on the EPPO A1 and A2 list, respectively (EPPO 2017a, b), but also for the National Priority Plant Pests in Australia (Plant Health Committee 2019) and other European Union regulated *Xanthomonas* plant pathogens. We envisage that our scheme will be useful to type more strains to further expand our understanding of the genetic structure and evolution, population expansions, and geographic dispersal of *X. citri*.

Acknowledgments

We thank Johanna Wong, Daniel J. E. McKnight, Fridtjof Snijders, John Webster, Paul Worden, and Gus McFarlane for their technical support provided during this study; and Bernie Dominiak and Krista Plett for reviewing a presubmission version of the manuscript.

Literature Cited

Alikhan, N.-F., Petty, N. K., Ben Zakour, N. L. B., and Beatson, S. A. 2011. BLAST Ring Image Generator (BRIG): Simple prokaryote genome comparisons. BMC Genomics 12:402.

Almeida, N. F., Yan, S., Cai, R., Clarke, C. R., Morris, C. E., Schaad, N. W., Schuenzel, E. L., Lacy, G. H., Sun, X., Jones, J. B., Castillo, J. A., Bull, C. T., Leman, S., Guttman, D. S., Setubal, J. C., and Vinatzer, B. A. 2010. PAMDB, a multilocus sequence typing and analysis database and website for plant-associated microbes. Phytopathology 100:208-215.

An, S.-Q., Potnis, N., Dow, M., Vorhölter, F.-J., He, Y.-Q., Becker, A., Teper, D., Li, Y., Wang, N., Bleris, L., and Tang, J.-L. 2020. Mechanistic insights into host adaptation, virulence and epidemiology of the phytopathogen *Xanthomonas*. FEMS Microbiol. Rev. 44:1-32.

Aritua, V., Harrison, J., Sapp, M., Buruchara, R., Smith, J., and Studholme, D. J. 2015. Genome sequencing reveals a new lineage associated with lablab bean and genetic exchange between *Xanthomonas axonopodis* pv. *phaseoli* and *Xanthomonas fuscans* subsp. *fuscans*. Front. Microbiol. 6:1080.

Bansal, K., Kumar, S., and Patil, P. B. 2022. Phylo-taxonogenomics supports revision of taxonomic status of 20 *Xanthomonas* pathovars to *Xanthomonas citri*. Phytopathology 112:1201-1207.

Bezdicek, M., Hanslikova, J., Nykrynova, M., Dufkova, K., Kocmanova, I., Kubackova, P., Mayer, J., and Lengerova, M. 2023. New multilocus sequence typing scheme for *Enterococcus faecium* based on whole genome sequencing data. Microbiol. Spectr. 11:e0510722.

Constantin, E. C., Cleenwerck, I., Maes, M., Baeyen, S., Van Malderghem, C., De Vos, P., and Cottyn, B. 2016. Genetic characterization of strains named as *Xanthomonas axonopodis* pv. *dieffenbachiae* leads to a taxonomic revision of the *X. axonopodis* species complex. Plant Pathol. 65:792-806.

Cunningham, C. W. 1997. Can three incongruence tests predict when data should be combined? Mol. Biol. Evol. 14:733-740.

da Gama, M. A. S., de Lima Ramos Mariano, R., da Silva Júnior, W. J., de Farias, A. R. G., Barbosa, M. A. G., da Silva Velloso Ferreira, M. A., Costa Júnior, C. R. L., Santos, L. A., and de Souza, E. B. 2018. Taxonomic repositioning of *Xanthomonas campestris* pv. *viticola* (Nayudu 1972) Dye 1978 as *Xanthomonas citri* pv. *viticola* (Nayudu 1972) Dye 1978 comb. nov and emendation of the description of *Xanthomonas citri* pv. *anacardii* to include pigmented isolates pathogenic to cashew plant. Phytopathology 108: 1143-1153.

Dubrow, Z. E., Carpenter, S. C. D., Carter, M. E., Grinage, A., Gris, C., Lauber, E., Butchachas, J., Jacobs, J. M., Smart, C. D., Tancos, M. A., Noël, L. D., and Bogdanove, A. J. 2022. Cruciferous weed isolates of *Xanthomonas campestris* yield insight into pathovar genomic relationships and genetic determinants of host and tissue specificity. Mol. Plant-Microbe Interact. 35:791-802.

EPPO. 2017a. A1 List of pests recommended for regulation as quarantine pest on 2020–09. https://www.eppo.int/ACTIVITIES/plant_quarantine/A1_list (accessed August 16, 2021).

EPPO. 2017b. A2 List of pests recommended for regulation as quarantine pest on 2021–09. https://www.eppo.int/ACTIVITIES/plant_quarantine/A2_list (accessed April 11, 2022).

Fargier, E., Saux, M. F.-L., and Manceau, C. 2011. A multilocus sequence analysis of *Xanthomonas campestris* reveals a complex structure within crucifer-attacking pathovars of this species. Syst. Appl. Microbiol. 34:156-165.

Ferraz, H. G. M., Badel, J. L., da Silva Guimarães, L. M., Reis, B. P., Tótola, M. R., Gonçalves, R. C., and Alfenas, A. C. 2018. *Xanthomonas axonopodis* pv. *eucalyptorum* pv. nov. causing bacterial leaf blight on eucalypt in Brazil. Plant Pathol. J. 34:269-285.

Gama, M. A. S., Mariano, R. L. R., Viana, F. M. P., Ferreira, M. A. S. V., and Souza, E. B. 2011. Polyphasic characterization of pigmented strains of *Xanthomonas* pathogenic to cashew trees. Plant Dis. 95:793-802.

Gevers, D., Cohan, F. M., Lawrence, J. G., Spratt, B. G., Coenye, T., Feil, E. J., Stackebrandt, E., Van de Peer, Y., Vandamme, P., Thompson, F. L., and Swings, J. 2005. Opinion: Re-evaluating prokaryotic species. Nat. Rev. Microbiol. 3:733-739.

Gordon, J. L., Lefeuvre, P., Escalon, A., Barbe, V., Cruveiller, S., Gagnevin, L., and Pruvost, O. 2015. Comparative genomics of 43 strains of *Xanthomonas citri* pv. *citri* reveals the evolutionary events giving rise to pathotypes with different host ranges. BMC Genomics 16:1098.

Graham, J. H., Gottwald, T. R., Cubero, J., and Achor, D. S. 2004. *Xanthomonas axonopodis* pv. *citri*: Factors affecting successful eradication of citrus canker. Mol. Plant Pathol. 5:1-15.

Hayward, A. C. 1993. The hosts of *Xanthomonas*. Pages 1-119 in: *Xanthomonas*. J. G. Swings and E. L. Civerolo, eds. Springer, Dordrecht, the Netherlands.

Hauben, L., Vauterin, L., Swings, J., and Moore, E. R. B. 1997. Comparison of 16S ribosomal DNA sequences of all *Xanthomonas* species. Int. J. Syst. Evol. Microbiol. 47:328-335.

Jacques, M.-A., Arlat, M., Boulanger, A., Boureau, T., Carrère, S., Cesbron, S., Chen, N. W. G., Cociancich, S., Darrasse, A., Denancé, N., Saux, M. F.-L., Gagnevin, L., Koebnik, R., Lauber, E., Noël, L. D., Pieretti, I., Portier, P., Pruvost, O., Rieux, A., Robène, I., Royer, M., Szurek, B., Verdier, V., and Vernière, C. 2016. Using ecology, physiology, and genomics to understand host specificity in *Xanthomonas*. Annu. Rev. Phytopathol. 54: 163-187.

Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., and Aluru, S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat. Commun. 9:5114.

Korber, B. 2000. HIV sequence signatures and similarities. Pages 55-72 in: Computational Analysis of HIV Molecular Sequences. A. G. Rodrigo and G. H. Learn, eds. Kluwer Academic Publishers, Dordrecht, the Netherlands.

Letunic, I., and Bork, P. 2021. Interactive Tree of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 49:W293-W296.

Lucena, L. P., Benko-Iseppon, A. M., Brenig, B., Azevedo, V., Aburjaile, F., Souza, E. B., and Gama, M. A. S. 2023. Draft genome sequence of seven pigmented strains of *Xanthomonas citri* pv. *anacardii*, the causal agent of cashew angular spot. Phytopathology 113:1360-1364.

Maiden, M. C. J. 2006. Multilocus sequence typing of bacteria. Annu. Rev. Microbiol. 60:561-588.

Maiden, M. C. J., Bygraves, J. A., Feil, E., Morelli, G., Russell, J. E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D. A., Feavers, I. M., Achtman, M., and Spratt, B. G. 1998. Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. Proc. Nat. Acad. Sci. U.S.A. 95:3140-3145.

Matsen, F. A., Mossel, E., and Steel, M. 2008. Mixed-up trees: The structure of phylogenetic mixtures. Bull. Math. Biol. 70:1115-1139.

Midha, S., and Patil, P. B. 2014. Genomic insights into the evolutionary origin of *Xanthomonas axonopodis* pv. *citri* and its ecological relatives. Appl. Environ. Microbiol. 80:6266-6279.

Monteiro, V. L. B., Melo, F. L., and Rossato, M. 2023. *Xanthomonas citri* infecting teak (*Tectona grandis*) in Brazil, characterization and copper resistance. Trop. Plant Pathol. 48:417-430.

Na, S.-I., Kim, Y. O., Yoon, S.-H., Ha, S.-m., Baek, I., and Chun, J. 2018. UBCG: Up-to-date bacterial core gene set and pipeline for phylogenomic tree reconstruction. J. Microbiol. 56:280-285.

Nei, M., and Gojobori, T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3:418-426.

Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T. G., Fookes, M., Falush, D., Keane, J. A., and Parkhill, J. 2015. Roary: Rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691-3693.

Parkinson, N., Aritua, V., Heeney, J., Cowie, C., Bew, J., and Stead, D. 2007. Phylogenetic analysis of *Xanthomonas* species by comparison of partial gyrase B gene sequences. Int. J. Syst. Evol. Microbiol. 57:2881-2887.

Parkinson, N., Cowie, C., Heeney, J., and Stead, D. 2009. Phylogenetic structure of *Xanthomonas* determined by comparison of *gyrB* sequences. Int. J. Syst. Evol. Microbiol. 59:264-274.

Payne, M., Kaur, S., Wang, Q., Hennessy, D., Luo, L., Octavia, S., Tanaka, M. M., Sintchenko, V., and Lan, R. 2020. Multilevel genome typing: Genomics guided scalable resolution typing of microbial pathogens. Eurosurveillance 5:1900519.

Pimentel, D., McNair, S., Janecka, J., Wightman, J., Simmonds, C., O'Connell, C., Wong, E., Russel, L., Zern, J., Aquino, T., and Tsomondo, T. 2001. Economic and environmental threats of alien plant, animal, and microbe invasions. Agric. Ecosyst. Environ. 84:1-20.

Plant Health Committee. 2019. National Priority Plant Pests (2019). https://www.agriculture.gov.au/biosecurity-trade/pests-diseases-weeds/plant/national-priority-plant-pests-2019 (accessed November 9, 2023).

Price, M. N., Dehal, P. S., and Arkin, A. P. 2010. FastTree 2–Approximately maximum-likelihood trees for large alignments. PLoS One 5:e9490.

Rademaker, J. L. W., Louws, F. J., Schultz, M. H., Rossbach, U., Vauterin, L., Swings, J., and de Bruijn, F. J. 2005. A comprehensive species to strain taxonomic framework for *Xanthomonas*. Phytopathology 95:1098-1111.

Ragupathy, R., Jolley, K. A., Zamuner, C., Jones, J. B., Redfern, J., Behlau, F., Ferreira, H., and Enright, M. C. 2023. Core-genome multilocus sequence typing for epidemiological and evolutionary analyses of phytopathogenic *Xanthomonas citri*. Appl. Env. Microbiol. 89:e0210122.

Ranjbar, R., Karami, A., Farshad, S., Giammanco, G. M., and Mammina, C. 2014. Typing methods used in the molecular epidemiology of microbial pathogens: A how-to guide. New Microbiol. 37:1-15.

Ruh, M., Briand, M., Bonneau, S., Darrasse, A., Jacques, M.-A., and Chen, N. W. G. 2021. Complete and circularized genome sequences of three *Xanthomonas* strains pathogenic on soybean and alfalfa. Microbiol. Resour. Announc. 10:e00537-21.

Sabat, A. J., Budimir, A., Nashev, D., Sá-Leão, R., van Dijl, J. M., Laurent, F., Grundmann, H., Friedrich, A. W., and ESCMID Study Group of Epidemiological Markers (ESGEM). 2013. Overview of molecular typing methods for outbreak detection and epidemiological surveillance. Eurosurveillance 18:20380.

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

Silva Junior, J. W., Farias, A. R. G., Lima, N. B., Benko-Iseppon, A. M., Aburjaile, F., Balbino, V. Q., Falcão, R. M., Leitão Paiva Júnior, S. S., Sousa-Paula, L. C., Mariano, R. L. R., Souza, E. B., and Gama, M. A. S. 2018. Complete genome sequence of *Xanthomonas citri* pv. *anacardii* strain IBSBF2579 from Brazil. Genome Announc. 6:e01574-17.

Simar, S. R., Hanson, B. M., and Arias, C. A. 2021. Techniques in bacterial strain typing: Past, present, and future. Curr. Opin. Infect. Dis. 34:339-345.

Sintchenko, V., Iredell, J. R., and Gilbert, G. L. 2007. Pathogen profiling for disease management and surveillance. Nat. Rev. Microbiol. 5:464-470.

Stamatakis, A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312-1313.

Starr, M. P. 1981. The genus *Xanthomonas*. Pages 742-763 in: The Prokaryotes: A Handbook on Habitats, Isolation and Identification of Bacteria. M. P. Starr, H. Stolp, H. G. Trüper, A. Balows, and H. G. Schlegel, eds. Springer-Verlag, Berlin, Germany.

Timilsina, S., Potnis, N., Newberry, E. A., Liyanapathiranage, P., Iruegas-Bocardo, F., White, F. F., Goss, E. M., and Jones, J. B. 2020. *Xanthomonas* diversity, virulence and plant–pathogen interactions. Nat. Rev. Microbiol. 18:415-427.

Tonkin-Hill, G., MacAlasdair, N., Ruis, C., Weimann, A., Horesh, G., Lees, J. A., Gladstone, R. A., Lo, S., Beaudoin, C., Floto, R. A., Frost, S. D. W., Corander, J., Bentley, S. D., and Parkhill, J. 2020. Producing polished prokaryotic pangenomes with the Panaroo pipeline. Genome Biol. 21:180.

Vauterin, L., Hoste, B., Kersters, K., and Swings, J. 1995. Reclassification of *Xanthomonas*. Int. J. Syst. Evol. Microbiol. 45:472-489.

Vurro, M., Bonciani, B., and Vannacci, G. 2010. Emerging infectious diseases of crop plants in developing countries: impact on agriculture and socio-economic consequences. Food Secur. 2:113-132.

Young, J. M., Park, D.-C., Shearman, H. M., and Fargier, E. 2008. A multilocus sequence analysis of the genus *Xanthomonas*. Syst. Appl. Microbiol. 31:366-377.

Zarei, S., Taghavi, S. M., Hamzehzarghani, H., Osdaghi, E., and Lamichhane, J. R. 2018. Epiphytic growth of *Xanthomonas arboricola* and *Xanthomonas citri* on non-host plants. Plant Pathol. 67:660-670.

Zeigler, D. R. 2003. Gene sequences useful for predicting relatedness of whole genomes in bacteria. Int. J. Syst. Evol. Microbiol. 53:1893-1900.

Zhou, Z., Alikhan, N.-F., Sergeant, M. J., Luhmann, N., Vaz, C., Francisco, A. P., Carriço, J. A., and Achtman, M. 2018. GrapeTree: Visualization of core genomic relationships among 100,000 bacterial pathogens. Genome Res. 28:1395-1404.

Zombre, C., Sankara, P., Ouédraogo, S. L., Wonni, I., Boyer, K., Boyer, C., Terville, M., Javegny, S., Allibert, A., Vernière, C., and Pruvost, O. 2016. Natural Infection of Cashew (*Anacardium occidentale*) by *Xanthomonas citri* pv. *mangiferaeindicae* in Burkina Faso. Plant Dis. 100: 718-723.