Phytopathology® • 2024 • 114:792-801 • https://doi.org/10.1094/PHYTO-09-23-0309-R

Population Biology

Diversity and Clonality in Populations of *Phytophthora citrophthora* and *P. syringae* Causing Brown Rot of Citrus in California

Nathan Riley,[1] Helga Förster,[2] and James E. Adaskaveg[1,†]

[1] Department of Plant Pathology, University of California, Riverside, CA 92521
[2] Department of Microbiology and Plant Pathology, University of California, Riverside, CA 92521
Accepted for publication 16 November 2023.

Abstract

Phytophthora citrophthora and *P. syringae* are currently the primary causal organisms of brown rot of citrus fruits in California. To possibly find an explanation for the prevalence of the previously minor species *P. syringae*, we determined the population structures of both pathogens in California using next-generation sequencing and population genomics analyses. Whole-genome sequencing and aligning with newly assembled reference genomes identified 972,266 variants in 132 isolates of *P. citrophthora* and 422,208 variants in 154 isolates (including 24 from noncitrus tree crops) of *P. syringae* originating from three major growing regions. The resulting data sets were visualized using principal component analysis, discriminant analysis of principal components, unweighted pair-group method with arithmetic mean dendrograms, fastStructure, and minimum spanning networks, and we obtained the index of association, diversity summary statistics, and genetic distance statistics values G_{ST}, G''_{ST}, and Jost's *D*. Subpopulations of both species were mostly defined by their geographic origin indicating restricted dispersal of inoculum. Except for five isolates, the population structure of *P. citrophthora* (that is heterothallic and unlikely to reproduce sexually) was clonal to semi-clonal, with very little genetic diversity within and among subgroups. In contrast, the population structure of *P. syringae* was also clonal to semi-clonal, but isolates were placed into four main clusters of much higher diversity. Clonality in both species can be explained by a high level of asexual reproduction. The higher diversity in the homothallic *P. syringae* is likely due to commonly occurring sexual reproduction. One distinct cluster of *P. syringae* consisted solely of isolates from noncitrus hosts; therefore, the origin of *P. syringae* in citrus could not be resolved.

Keywords: homothallic heterothallic, population diversity, population genomics, population structure

Citrus is an important crop in many countries with subtropical and tropical environments. Total worldwide production was estimated at 143,755,600 metric tons of fruit in 2020 (FAO 2021). In California, citrus production was valued at $2.5 billion in 2021/22 (Anonymous 2022). The California citrus industry expanded from a first commercial orchard in Los Angeles in 1841 to other southern portions of the state. Currently, most of the citrus production is further North in the Central Valley (CV), particularly in Fresno, Kern, and Tulare counties. The remaining production is primarily in the original growing areas of Southern California in the Inland Empire (IE; i.e., Riverside and San Bernardino counties), as well in coastal Ventura County (VE) and San Diego County (Geisseler and Horwath 2016).

Species of *Phytophthora* are responsible for various diseases affecting citrus trees throughout the growing season. These include root rot of young seedlings and mature trees, foot rot causing gummosis of the trunk (i.e., trunk cankers), and brown rot of fruit, all of which may be caused by some of the same species (Graham and Feichtenberger 2015). Fruit brown rot is characterized by tan to brown, leathery lesions that have a distinctive odor. Infected fruit remain firm unless secondary infections occur. In California, brown rot can be caused by *P. citrophthora*, *P. hibernalis*, *P. nicotianae*, or *P. syringae*. *P. syringae* and *P. hibernalis* have been considered of minor importance in the past (Feld et al. 1979; Graham and Menge 2000; Klotz 1978). More recently, however, *P. citrophthora* and *P. syringae* were demonstrated to be the major pathogens, and they can occur within the same orchard (Hao et al. 2018). In other citrus-producing regions, such as Florida, additional species such as *P. palmivora* and *P. citricola* have been found responsible for brown rot (Graham and Menge 2000).

The shift in species composition of brown rot pathogens in California has caused trade restrictions for the California citrus industry by China, a major export country where *P. hibernalis* and *P. syringae* are quarantine pathogens (Hao et al. 2018). Export of California citrus fruit to China was closed from 2013 to 2014 and again in 2015 due to the detection of *P. syringae* and *P. hibernalis* in orange fruit shipments, resulting in economic loss to growers, packinghouses, and shippers (Adaskaveg and Förster 2014). The status of *P. syringae* and *P. hibernalis* as quarantine pathogens has elevated brown rot of citrus from a troublesome inconvenience to a serious problem over the past decade.

More information is needed as to why *P. syringae* has become a dominant brown rot pathogen in California, and thus, our goal was to study the population demographics of this species in comparison with *P. citrophthora*. The two species differ ecologically: *P. citrophthora* can be isolated year-round, whereas *P. syringae* is only active during the cooler months. Furthermore, *P. syringae* is homothallic and capable of self-crossing, whereas *P. citrophthora* is heterothallic but not thought to reproduce in the environment (Erwin and Ribeiro 1996). These characteristics may influence population structure and diversity. Population genomic analyses were recently utilized on other *Phytophthora* species such as *P. pluvialis* (Tabima et al. 2021), *P. capsici* (Vogel et al. 2021), and *P. infestans* (Zhang et al. 2021) to investigate population structure, fungicide sensitivity, and pathogen adaptation. Our specific objectives were to (i) collect isolates of *P. syringae* and *P. citrophthora* from citrus production

†Corresponding author: J. E. Adaskaveg; jim.adaskaveg@ucr.edu

Funding: Support was provided by the California Citrus Research Board.

e-Xtra: Supplementary material is available online.

The author(s) declare no conflict of interest.

areas of California not previously surveyed by Hao et al. (2018); (ii) characterize the population structure of these species and determine if populations sampled from different geographic regions are genetically distinct; and (iii) determine if isolates of *P. syringae* from noncitrus hosts are genetically distinct from those collected from citrus. To achieve these goals, reference isolates of *P. syringae* and *P. citrophthora* were sequenced, and sequences were assembled into reference genomes. Subsequently, 132 isolates of *P. citrophthora* and 154 isolates of *P. syringae* from major growing regions in California were sequenced, and sequences were aligned to the reference genomes to identify variant (i.e., single-nucleotide polymorphism; SNP) markers to be used for population genomic analyses.

Materials and Methods

Isolate collection and variant (SNP) identification

Isolates of *P. citrophthora* and *P. syringae* collected from brown-rotted fruit in the CV of California between 2013 and 2015 were used from a previous study where the species identity of isolates was verified by morphology, random amplified polymorphic DNA fragment patterns, and using a multiplex TaqMan qPCR assay with species-specific probes (Hao et al. 2018). Additional isolates were obtained from VE and the IE region from symptomatic citrus fruit, leaf litter, and soil, as well as from the soil of noncitrus hosts between 2018 and 2022. Isolations were done as described by Hao et al. (2018). Colonies resembling *Phytophthora* spp. were subcultured onto clarified 5% V8 agar (Ribeiro 1978). Species identification of newly collected isolates was based on morphology, and a subset was subjected to TaqMan qPCR using species-specific probes. For genome sequencing, a total of 132 isolates of *P. citrophthora* (70 from the CV, 36 from VE, and 26 from the IE) and 130 isolates of *P. syringae* (80 from the CV, 14 from VE, and 36 from the IE) were selected from citrus sources. Additionally, 24 isolates of *P. syringae* were included from orchard soils of noncitrus hosts (19 from almond, 4 from walnut, and 1 from cherry). The isolates used in this study are listed in Table 1.

For DNA extraction, isolates were grown for 5 to 10 days in 10% clarified V8 broth at 20°C. Mycelia were rinsed three times with sterile water and lyophilized. Tissues were homogenized using a stainless-steel grinding ball (Steelball LysingMatrix; MP Biomedicals, LLC, Irvine, CA) using the FastPrep-24 instrument (MP Biomedicals, LLC) at 6.0 m/s twice for 45 s, and DNA was extracted using the DNeasy Plant Mini Kit (Qiagen, Germantown, MD). DNA quality was assessed by measuring the 260/280 nm ratio of absorbance (NanoDrop ND-1000 UV-Vis Spectrophotometer; Thermo Fisher Scientific, Waltham, MA) and by gel electrophoresis in a 2% agarose gel. Prior to the sequencing of the selected isolates, reference genomes were constructed for *P. citrophthora* isolate 2440 and *P. syringae* isolate 4571. DNA was quantified (Qubit 4 Fluorometer; Invitrogen, Carlsbad, CA) before submission for library prep at the University of California Riverside Genomics Core Facility and subsequent 150-bp paired-end sequencing on the Illumina NovaSeq platform. Processing of raw data, creation of reference assemblies, alignment of population data, variant calling, and variant filtering were performed on the UCR High Performance Computing Center cluster. Raw reads were filtered for a minimum quality score of 30, and Illumina adapters were removed using the fastq-mcf tool (ea-utils v.1.1.2-537; Aronesty 2013). Filtered raw reads were then used for reference genome assembly using the SPAdes assembly program v.3.13.1with default settings (Bankevich et al. 2012). For quality assessment, reference assemblies were analyzed for basic characteristics and quality using QUAST v.4.6.3 (Gurevich et al. 2013). Assemblies were then assessed for completeness using BUSCO v.3.0.2 in the "genome" mode with the "Eukaryota Odb9" database (Manni et al. 2021). DNA from the selected isolates was extracted as described above and submitted for library preparation and 150-bp paired-end sequencing on the Illumina NovaSeq platform. Initial processing of data was performed as described above. The reference assemblies for *P. citrophthora* and *P. syringae*, as well as the raw data for all other isolates sequenced, have been deposited with links to BioProject accession number PRJNA1010852 in the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/bioproject/). The reads from the sequenced isolates were mapped to their respective reference genomes. Using BWA v.0.7.17 (Li 2013), reference assemblies were indexed using default settings, and alignments were made using the MEM algorithm with the -M setting activated. All other settings were in their default modes. The resulting BAM files were then sorted, mate-fixed, and marked for duplicates

TABLE 1. Isolates of *Phytophthora* spp. from California used in this study and their host, geographic origin, and year of isolation

Species	Host	Geographic origin	Year	Number of isolates
P. citrophthora	*Citrus reticulata*, fruit	Central Valley	2014	1
	C. sinensis, fruit	Central Valley	2012	1
	C. sinensis, fruit	Central Valley	2013	4
	C. sinensis, fruit	Central Valley	2014	53
	C. sinensis, fruit	Central Valley	2016	4
	C. sinensis, fruit	Central Valley	2017	1
	C. sinensis, fruit	Inland Empire	2018	1
	C. sinensis, fruit	Inland Empire	2019	25
	C. sinensis, root	Ventura Co.	2000[a]	1
	C. sinensis, fruit	Ventura Co.	2019	36
	C. sinensis, soil	Central Valley	2016	6
P. syringae	*C. limon*, fruit	Central Valley	2014	1
	C. sinensis, fruit	Central Valley	2012[a]	11
	C. sinensis, fruit	Central Valley	2013	18
	C. sinensis, fruit	Central Valley	2014	41
	C. sinensis, fruit	Central Valley	2016	2
	C. sinensis, fruit	Central Valley	2017	1
	C. sinensis, fruit	Inland Empire	2019	36
	C. sinensis, fruit	Ventura Co.	2019	14
	C. sinensis, leaf litter	Central Valley	2016	6
	C. sinensis, leaf litter	Central Valley	2017	1
	Juglans regia, soil	Central Valley	2017	4
	Prunus avium, soil	Central Valley	2020	1
	P. dulcis	Central Valley	2007	2
	P. dulcis	Central Valley	2020	3
	P. dulcis, soil	Central Valley	2017	14

[a] Isolates *P. syringae* 2440 and *P. citrophthora* 4571 were used for construction of reference genomes.

using sort, fixmate, and markdup functions (all with default settings except for fixmate, which utilized -m to add mate score tags for use by markdup) of SAMtools v.1.14 (Danecek et al. 2021; Li et al. 2009). SAMtools was also used to assess mapping percentages with the flagstat function. Genomic variants were predicted using GATK v.4.2.5.0 (McKenna et al. 2010). For both species, each isolate was independently genotyped with GATK's haplotype caller in the GVCF mode. Variants were validated using the ValidateVariants function before being pooled using the CombineGVCF function (both using default settings), and the combined GVCF file was then genotyped using the GenotypeGVCFs function, also with default settings. A final combined GVCF file was created for each species using the GATK SelectVariants function with –restrict-alleles-to set to "BIALLELIC", –remove-unused-alternates set to "true", –select-typeto-include set to "SNP", and –exclude-non-variants set to "true". Variants were then filtered using VCFTools v.0.1.16-18 (Danecek et al. 2011), in which all singletons, all variants with more than 10% missing data, and all variants with a read depth of less than 4 and a quality score of less than 10 were removed. Following this, RepeatModeler v.2.0.4 (Flynn et al. 2020) was utilized to identify repetitious regions in the genome (using the BuildDatabase and RepeatModeler functions), and RepeatMasker v.4-1-4 (Tarailo-Graovac and Chen 2009) was used to perform soft masking and generate a ".gff" file that was then converted into a ".bed" file containing that information. This ".bed" file was then used to filter out variants located in those repetitious regions with the –exclude-bed setting of VCFtools.

Population structure

All analyses were conducted in R v.4.1.3 (R Core Team 2022) on non-clone-corrected data unless otherwise specified. Principal components analysis (PCA) and discriminant analysis of principal components (DAPC) were conducted for *P. citrophthora* and *P. syringae* using the *adegenet* package v.2.1.10 (Jombart 2008) with nf = 4 and nf = 5, respectively. PCA scores were plotted using the *ggplot2* package v.3.4.2 (Wickham 2016). DAPC scatterplots were plotted using the scatter.dapc function of *adegenet*, with n.pca = 50 and n.da = 3 for *P. citrophthora* and n.pca = 20 and n.da = 4 for *P. syringae*. Isolates of both species were grouped by geographic region. Isolates for *P. syringae* were also plotted by host of origin. Unweighted pair-group method with arithmetic mean (UPGMA) dendrograms for each species were generated based on bitwise distance using the aboot function of the *poppr* package v.2.9.4 (Kamvar et al. 2014) with 1,000 permutations. These were then plotted using the packages *ggtree* v.3.7.2 and *ggbreak* v.0.1.1 to extend the capabilities of ggplot2 (Xu et al. 2021, 2022). The gl2structure function from the R package *dartR* v.2.7.2 (Gruber et al. 2018) was used to create structure files for use in population structure analyses. The Python program fastStructure v.1.0 (Raj et al. 2014) was then used to analyze the population structure for each species, and the structure.py function was executed for $K = 1$ to 10. The chooseK.py function was used to identify K values that best explained the structure in our data, and $K = 2$ and $K = 4$ were found to be the best fit for *P. citrophthora* and *P. syringae*, respectively. Unaligned results for K1-10 are shown for each species in Supplementary Figure S1. Population structure data for those runs were then plotted using the *pophelper* package v.2.3.1 in R (Francis 2017), and the results were aligned to the UPGMA dendrograms generated for each species.

Clonality and minimum spanning networks (MSNs)

To infer the level of clonality in the sampled populations, the index of association (I_A) was calculated as described in Shakya et al. (2021) with modifications. The following simulations and actual calculations were performed for each species overall, as well as for subsets containing only citrus isolates from each regional population and for the noncitrus isolates of *P. syringae*. First, the glSim function of the *adegenet* v.2.1.10 package was used to simulate clonal (90% linked SNPs), partially clonal (50% linked SNPs), and sexual (10% linked SNPs) populations based on each data set. Then, the samp.ia function from the *poppr* v.2.9.4 package (Kamvar et al. 2014) was used to calculate the I_A for each data set and their simulations from 10,000 randomly sampled SNPs with 100 replicates. Significant differences were identified using analysis of variance followed by Tukey's honestly significant difference test in R. Due to the presence of clonality, the *poppr* package was then used to generate color-coded MSN plots with distance filters of 0.00375 and 0.04 for *P. citrophthora* and *P. syringae*, respectively, based on bit-wise distances observed in the UPGMA analyses. A second color-coded MSN plot was generated for *P. syringae* to reflect the host of the isolates.

Differentiation statistics

The utils.basic.stats function of the R package *dartR* v.2.7.2 (Gruber et al. 2018) was used to generate basic summary statistics, including observed heterozygosity (H_O), within-population gene diversity (H_S), total genetic diversity (H_T), gene diversity among samples (D_{ST}), and fixation index (F_{ST}, D_{ST}/H_T). The pairwise Gst_Nei, pairwise_Gst_Hedrick, and pairwise_D functions from the R package *mmod* v.1.3.3 (Winter 2012) were used to calculate Nei's G_{ST} (a multi-allelic measure of F_{ST}) (Nei 1973; Nei and Chesser 1983), Hedrick's G''_{ST} (a version of G_{ST} that accounts for sample size) (Meirmans and Hedrick 2011), and Jost's D (a measure of allele frequency difference among populations), respectively (Jost 2008).

Results

Variant (SNP) identification

The quality and structure metrics of reference genome assemblies for *P. citrophthora* and *P. syringae*, including size, number of contigs, N50, L50, and error rate, are listed in Table 2. BUSCO outputs for the genome assemblies of both species can be found in Table 3. The results show the proportions of orthologous genes to the 303 genes that were searched for as a metric for assem-

TABLE 3. BUSCO output for reference genome assemblies of *Phytophthora citrophthora* and *P. syringae*

	Species			
	P. citrophthora		*P. syringae*	
BUSCO category	% of total	*N*	% of total	*N*
Complete	92.8%	281	92.4%	280
Complete, single copy	89.8%	272	89.1%	270
Complete, duplicated	3.0%	9	3.3%	10
Fragmented	1.7%	5	1.3%	4
Missing	5.5%	17	6.3%	19
Total	–	303	–	303

TABLE 2. QUAST output for reference genome assemblies of *Phytophthora citrophthora* and *P. syringae*[a]

Species	Number of contigs[b]	Largest contig	Total length	N50	L50	GC%	*N*/100 kbp
P. citrophthora	4,904	1,537,029	42,171,630	37,026	274	51.8	274.58
P. syringae	6,950	153,696	52,547,088	22,629	629	52.8	27.58

[a] N50 is the minimum contig length making up at least 50% of the assembly; L50 is the number of contigs that make up 50% of the assembly; GC% is the percentage of G and C nucleotides within the assembly; and *N*/100 kbp is the average number of uncalled bases per 100,000 bases in the assembly.
[b] Contigs of >500 bp.

bly completeness. Reads of all sequenced isolates of both species aligned more than 90% to their respective reference genomes. A total of 945,881 variants were identified in *P. syringae* before filtering and repeat masking, whereas 1,531,374 variants were identified in *P. citrophthora*. After filtering out singletons and variants with more than 10% missing data, a read depth of less than 4, or a quality score of less than 10, as well as singletons and variants occurring in repeated regions, the final data set for *P. syringae* contained 422,208 variants, whereas the final data set for *P. citrophthora* contained 972,266 variants.

Population structure

The PCA for isolates of *P. citrophthora* from citrus by geographic region of origin resulted in three distinct clusters, and 85.51% of the observed variance was captured by PC1 (Fig. 1A). Most isolates were placed into two large closely related clusters to the right of the *y* axis, whereas five CV isolates were found more distant to the left of the *y* axis. The large cluster above the *x* axis primarily consisted of VE and IE isolates, whereas the one below the *x* axis primarily consisted of CV isolates. Still, both clusters contained isolates from all three geographic regions.

The PCA for *P. syringae* isolates grouped by geographic origin revealed three clusters, and 47.07% of the observed variance was captured by PC1 (Fig. 1B). A large cluster in the lower left of the plot consisted almost entirely of citrus isolates from the CV and IE, except for three isolates from VE. A smaller cluster in the lower right comprised 21 noncitrus CV isolates. Eleven citrus VE isolates were loosely clustered along the *y* axis in the upper left of the plot, whereas one citrus CV and three almond CV isolates were not clearly associated with any cluster.

In the DAPCs where differences between groups are maximized and similarities are minimized, isolates of the two main PCA clusters for *P. citrophthora* were separated into three distinct clusters in the scatterplot (Fig. 1C). Each of these three clusters contained isolates mostly from one of the regional groups (CV, IE, or VE), except for one VE isolate within the CV cluster. The DAPC scatterplot of *P. syringae* based on the geographic origin of isolates also formed distinct clusters for each regional group, but the CV isolates grouped closely to IE isolates and several CV individuals plotted within the IE cluster. The cluster of 12 VE isolates was more distant, indicating higher differentiation of this population (Fig. 1D). One individual from VE was plotted within the IE cluster and a second

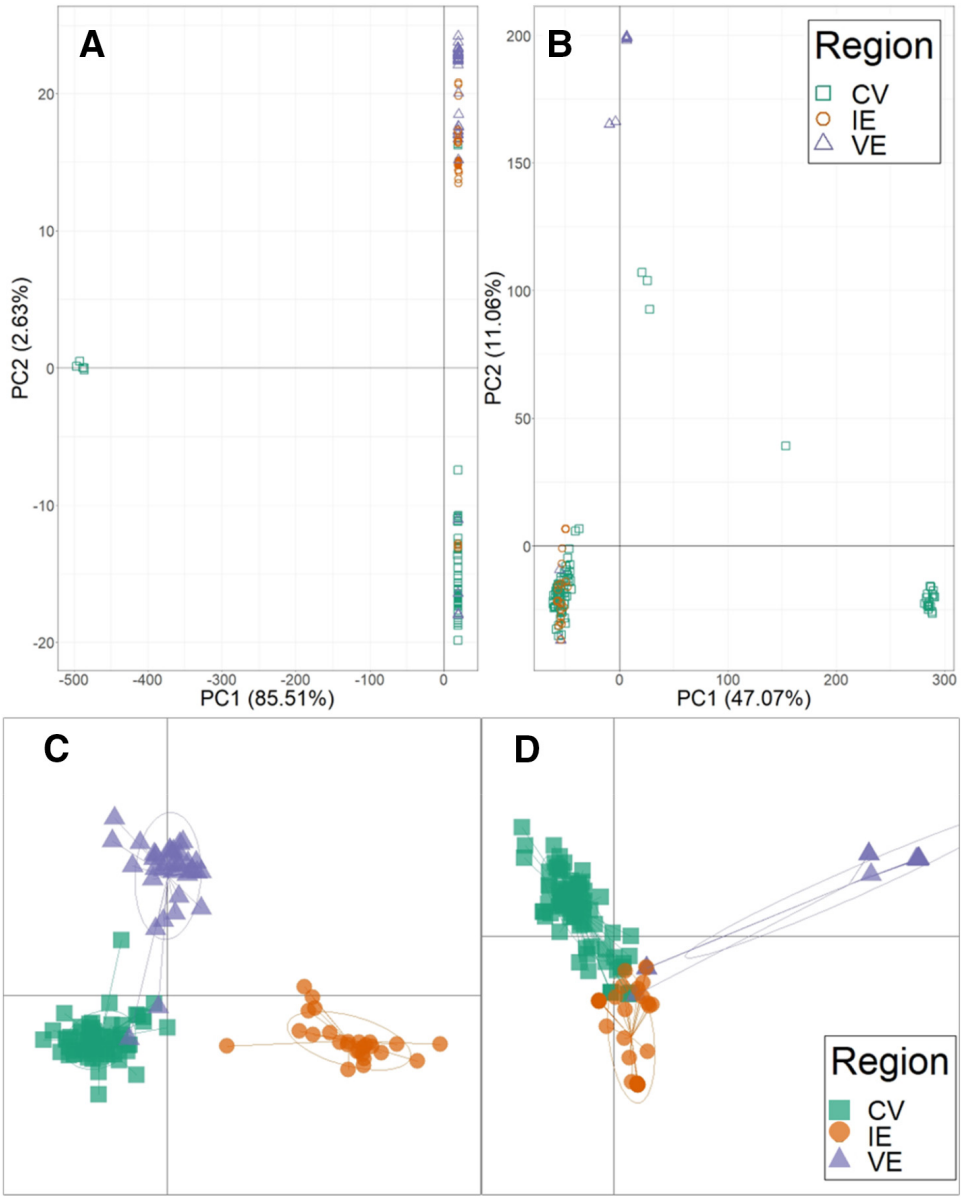

Fig. 1. A and B, Principal component (PC) analyses and **C and D,** discriminant analyses of PC visualizing the observed variance of A and C, *Phytophthora citrophthora* and B and D, *P. syringae* isolates grouped by geographic region. CV = Central Valley; IE = Inland Empire; and VE = Ventura County.

one (hidden behind the orange triangles of IE isolates in Fig. 1D) just outside of it.

In the second PCA for *P. syringae* isolates with categorization by host, the large cluster in the lower left of the plot and the looser cluster of 11 isolates in the upper left both consisted entirely of isolates from citrus fruit or leaf litter (Fig. 2A). A cluster of 21 isolates in the lower right originated entirely from noncitrus hosts (i.e., almond, cherry, and walnut). Three isolates from almond and one citrus CV isolate were not associated with any cluster.

The DAPC scatterplot of *P. syringae* with grouping by host reflects the distinctness of citrus isolates (except for one) from those collected from noncitrus hosts (Fig. 2B; green circles representing citrus leaf litter isolates are hidden behind the orange circles of the citrus fruit isolates). The 19 almond isolates formed a less tight cluster, and one isolate from walnut was closely associated. The remaining noncitrus isolates (three from walnut, one from cherry) were scattered more distantly from the almond isolates.

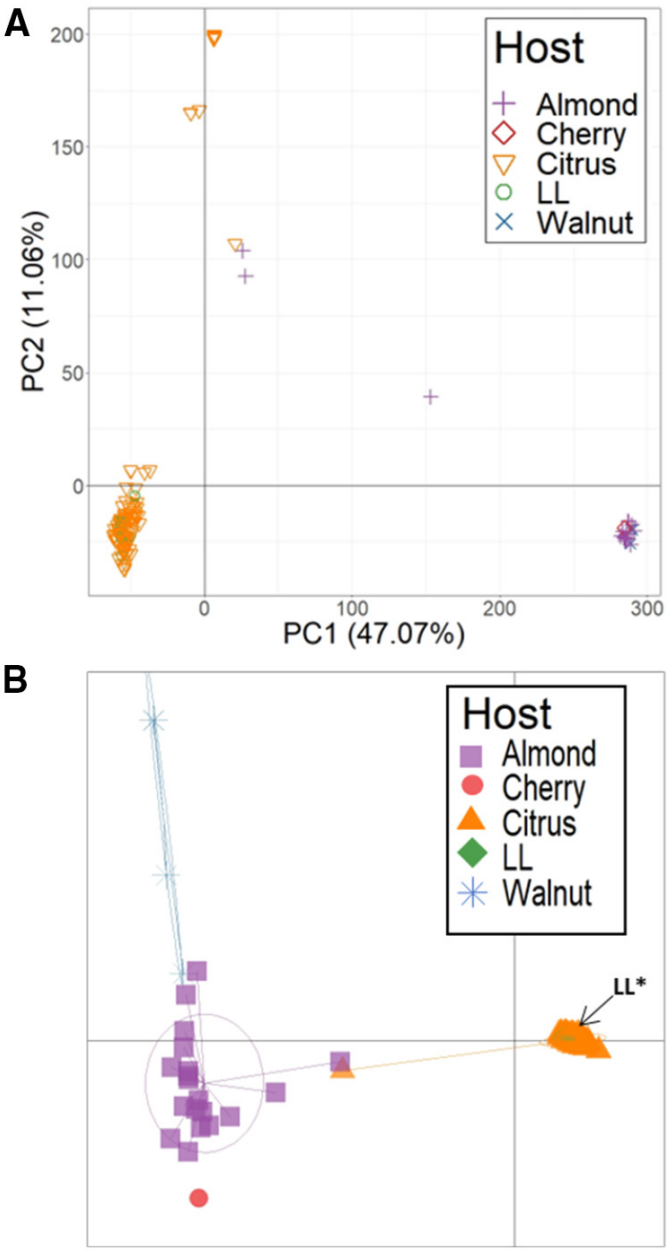

Fig. 2. A, Principal component (PC) analysis and **B**, discriminant analysis of PC to visualize observed variance of *Phytophthora syringae* isolates grouped by host. LL = citrus leaf litter (LL* = all LL isolates are within this cluster).

UPGMA dendrograms of *P. citrophthora* (Fig. 3A) and *P. syringae* (Fig. 3B) provided additional information on the population structures of the two species. Isolates of *P. citrophthora* exhibited relatively low genetic distances among individuals throughout the geographic populations sampled, except for five CV isolates (at the top of the dendrogram). The genetic diversity among these five isolates was much higher than that among the remaining majority of CV, IE, and VE isolates that were mostly found within distinct portions of the dendrogram and, thus, grouped based on their geographic origin. CV isolates were distinct, were located on a major lower branch of the dendrogram, and were separated from IE and VE isolates. In comparison, the *P. syringae* dendrogram showed much greater genetic distances among individuals, both within and between sampled geographic populations. The noncitrus isolates formed their own distinct cluster (at the top of the dendrogram in Figure 3B) separate from the other sampled populations. Isolates from VE clustered mostly on a branch in the upper parts of the dendrogram, whereas CV and IE isolates were not clearly separated.

When UPGMA dendrograms were aligned with fastStructure bar plots, similar groupings of isolates of both species were observed in the clusters generated by both types of analyses. Thus, for *P. citrophthora* most isolates grouped together (cluster 2 in Figure 3A). The exception again was five CV isolates that were assigned to their own cluster (cluster 1). This separation was observed for a range of K values from 2 to 10, indicating no mixture with the other populations. For *P. syringae*, isolates from noncitrus hosts again formed a distinct cluster (cluster 3 in Figure 3B). Similar to the UPGMA dendrogram, VE isolates from citrus were primarily assigned to their own unique cluster (cluster 2). Most CV and IE citrus isolates were assigned to two distinct clusters (clusters 4 and 1, respectively), although many isolates showed partial assignment to both clusters.

Clonality and MSN

Calculated I_A values across all loci indicated partial clonality ($I_A > 0$) in all sampled geographic populations of both species (Table 4). Replicates for each data set had very low deviation, and Tukey's honestly significant difference test found all actual data sets to be significantly different ($P < 0.05$) from the simulated data sets. For *P. citrophthora*, the overall and CV data sets had I_A values much greater than that of the simulated clonal data set, indicative of clonal reproduction, whereas the IE and VE data sets had I_A values closest to those of the semi-clonal datasets. For *P. syringae*, the overall and VE datasets had I_A values greater than that of the simulated clonal data set, whereas the IE and noncitrus data sets had I_A values closest to that of the simulated clonal data set. The CV data set had the lowest I_A value, and this was closest to that of the simulated semiclonal data set.

TABLE 4. Average index of association (I_A) for simulated and actual data sets for overall and regional populations of *Phytophthora citrophthora* and *P. syringae*

		I_A values[b]	
Data evaluation[a]	Data set	*P. citrophthora*	*P. syringae*
Simulated	Sexual	0.00329	0.00331
	Semiclonal	0.0645	0.0645
	Clonal	0.168	0.167
Actual	Overall	0.842	0.231
	Central Valley	0.889	0.0557
	Inland Empire	0.0332	0.135
	Ventura Co.	0.0325	0.418
	Noncitrus[c]	–	0.149

[a] Simulated data sets were calculated from 10,000 randomly selected single-nucleotide polymorphisms for 100 replicates.
[b] For each species, all actual data sets were significantly different from simulated data sets based on an analysis of variance and Tukey's honestly significant difference mean separation procedures ($P < 0.05$).
[c] All isolates from this data set originated from noncitrus hosts, whereas all other data sets had only isolates from citrus.

After distance filtering, the 132 *P. citrophthora* isolates and 154 *P. syringae* isolates were assigned to 38 and 70 multilocus genotypes (MLGs), respectively. The generated MSN for *P. citrophthora* revealed three large MLGs and smaller MLGs, each primarily consisting of isolates from one of the three sampled regions (Fig. 4A). The thick and dark connecting lines between nodes indicate small genetic distances, as shown by the scale bar in the figure. The MSN for *P. syringae* was comparatively more dispersed and included more nodes and a larger number of mostly smaller MLGs (Fig. 4B). One of the larger MLGs consisted of only IE isolates, whereas the other two were mostly associated with IE isolates but also were represented at relatively high incidence by one or two of the other geographic populations. The smaller groups of MLGs were mostly associated with CV isolates, although three represented isolates from VE, and they were connected by more gray and thinner lines, indicating larger genetic distances based on the scale bar. The genetic distances between the MLGs of CV isolates at the "tail" trailing off the top of the network were greater than in the rest of the network. When isolates of *P. syringae* were color-coded by host, it was evident that the nodes of the "tail" all represented noncitrus host isolates, and the remaining network all comprised isolates from citrus (Fig. 5). All except four of the 24 noncitrus isolates possessed their own unique MLG, and two each of these four originated from the same orchards.

Differentiation statistics

The summary statistics for *P. citrophthora* indicate relatively low within-population diversity H_S that was highest for the CV population (Table 5). The summary statistics for *P. syringae* indicate higher overall within-population diversity H_S and F_{ST} fixation index values as compared with *P. citrophthora*; for citrus isolates, H_S was highest for the VE population but was even higher for the noncitrus isolates. Analysis of pairwise Nei's G_{ST}, Hedrick's G''_{ST}, and Jost's D provided additional insight into the differentiation of populations of

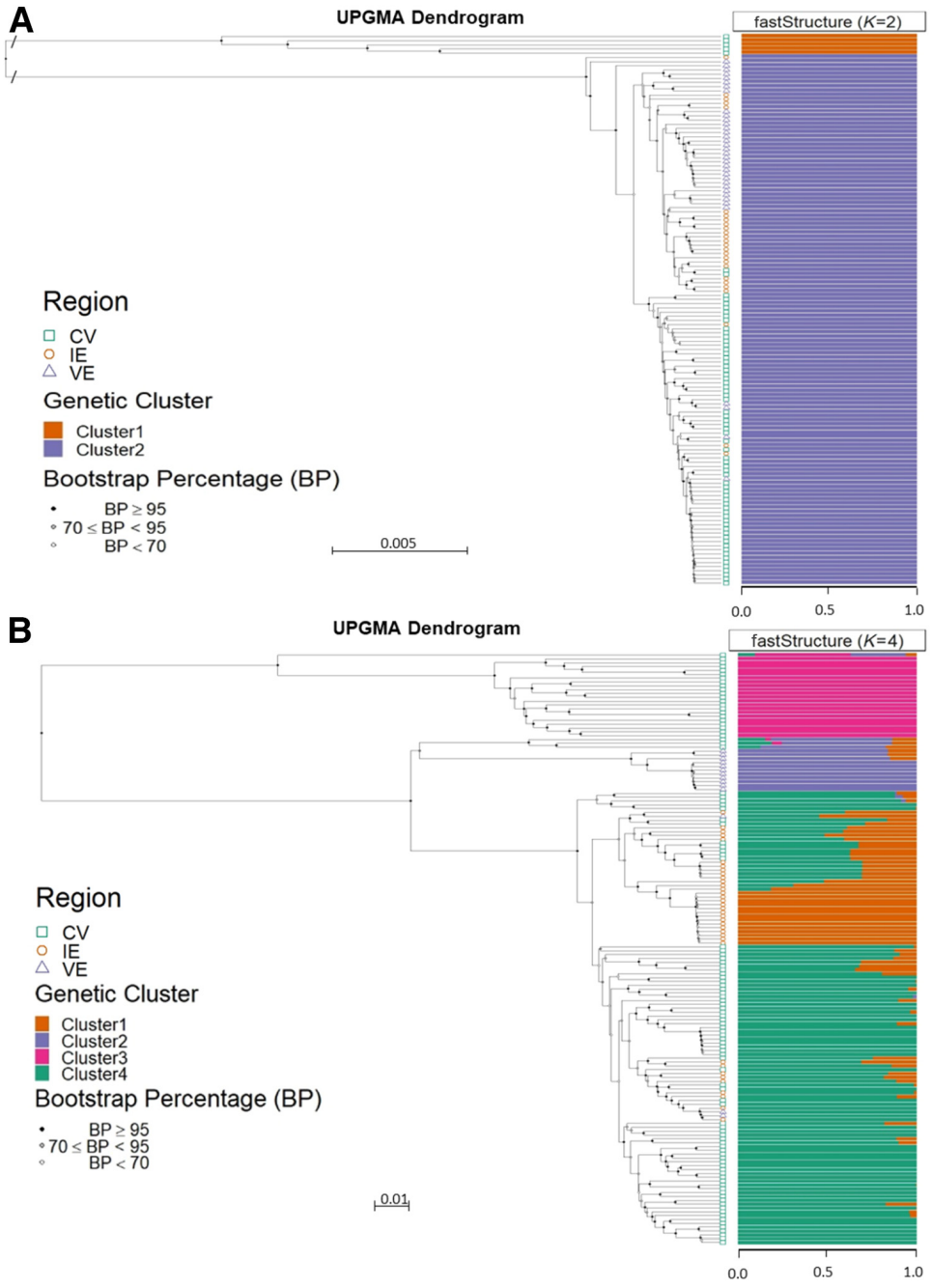

Fig. 3. Population structure of **A,** *Phytophthora citrophthora* and **B,** *P. syringae* visualized in unweighted pair-group method with arithmetic mean (UPGMA) dendrograms that were constructed from observed variants and aligned to bar plots of the fastStructure output for each isolate. In the legend, the color codes of the symbols refer to the geographic origin of the isolates (CV = Central Valley; IE = Inland Empire; and VE = Ventura Co.). In the fastStructure graphs, colors within a given bar represent the posterior membership probability of the isolate to be assigned to the predicted clusters that are color-coded in the squares of the legend.

P. citrophthora and *P. syringae* (Table 6). *P. citrophthora* exhibited lower levels of differentiation based on all metrics than *P. syringae*. For *P. citrophthora*, differentiation was lowest between VE and IE populations and highest between VE and CV populations. For isolates of *P. syringae* from citrus, differentiation was lowest between CV and IE populations and highest between VE and IE populations. Differentiation between isolates of *P. syringae* from noncitrus hosts and any of the three regional citrus populations was all higher than among the citrus populations.

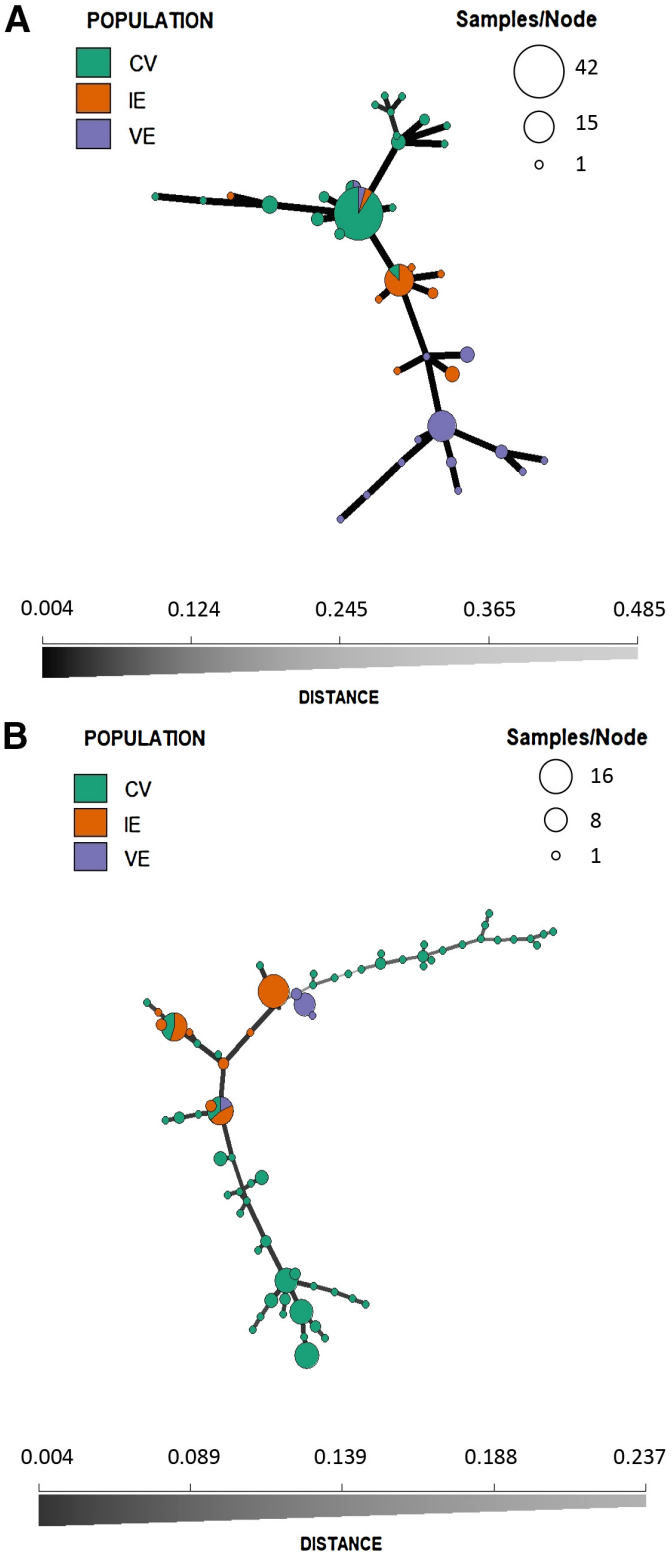

Fig. 4. Minimum spanning networks visualizing relationships of **A,** *Phytophthora citrophthora* and **B,** *P. syringae* isolates grouped by region (CV = Central Valley, IE = Inland Empire; and VE = Ventura Co.). Nodes represent individual multilocus genotypes (MLGs). The distance scale represents bitwise distance (connecting lines) over all loci in an MLG. The size differences of the circles represent the number of individuals within that MLG.

Discussion

Extensive population genetic analyses were conducted to characterize isolates of *P. citrophthora* and *P. syringae* from citrus in three major growing regions of California: the CV, VE, and the IE. New insights into population structures were gained, and we characterized distinct differences between the two brown rot pathogens that relate to their lifestyles. Isolates of *P. syringae* from noncitrus tree

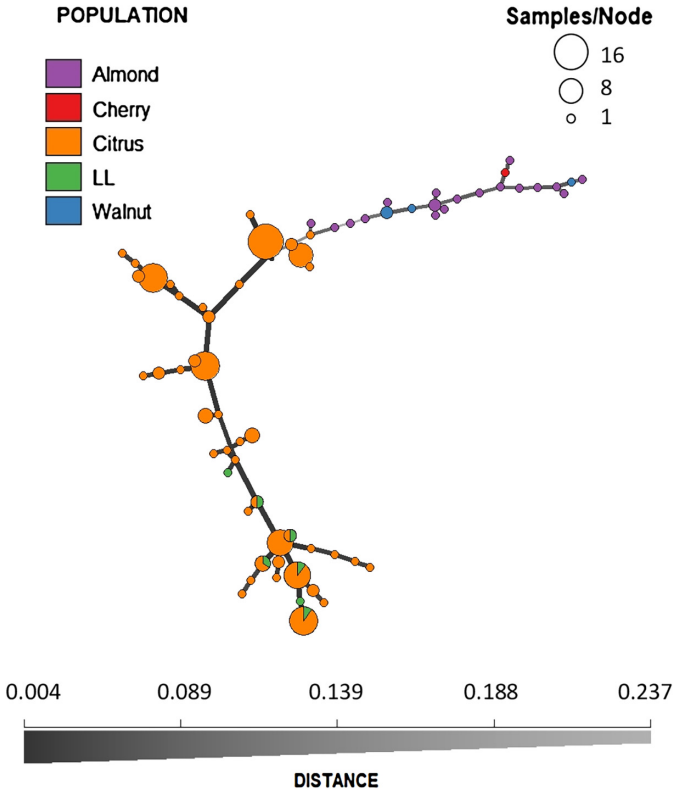

Fig. 5. Minimum spanning network visualizing relationships of *Phytophthora syringae* grouped by host type. Nodes represent individual multilocus genotypes (MLGs). The distance scale represents bitwise distance (connecting lines) over all loci in an MLG. The size differences in the circles represent the number of individuals within that MLG.

TABLE 5. Summary statistics for the genetic diversity of the overall population and geographic subpopulations of *Phytophthora citrophthora* and *P. syringae*

Species	Population	Measures of diversity[a]				
		H_O	H_S	H_T	D_{ST}	F_{ST}
P. citrophthora	Overall	0.1103	0.0700	0.0709	0.0009	0.012
	Central Valley	0.1475	0.1127	–	–	–
	Ventura Co.	0.0914	0.0484	–	–	–
	Inland Empire	0.0920	0.0485	–	–	–
P. syringae[b]	Overall	0.1125	0.1301	0.2374	0.1074	0.4522
	Central Valley	0.0841	0.0928	–	–	–
	Ventura Co.	0.1066	0.1212	–	–	–
	Inland Empire	0.0921	0.0856	–	–	–
	Noncitrus	0.1670	0.2218	–	–	–

[a] Measures of diversity are as follows: H_O: observed heterozygosity; H_S: within-population gene diversity; H_T: total genetic diversity; D_{ST}: gene diversity among populations; and F_{ST}: and fixation index, D_{ST}/H_T.
[b] Only isolates from citrus were included for the overall, Central Valley, Ventura Co., and Inland Empire populations of *P. syringae*.

crops separated from citrus populations, and, therefore, we found no evidence for the origin of this species on citrus and its emergence as a primary citrus brown rot pathogen.

Genomic variants in isolates were identified by aligning short-read sequences with our newly assembled reference genomes of the two species. No reference genomes were available at the time these studies were initiated. Recently published genomes of *P. syringae* with assembly sizes of 74.9 or 57 Mb (GenBank accessions 19064958 and 35224038), however, are of a similar size range as our genome of this species (i.e., 52 Mb). We conducted no genetic comparisons of these genomes including gene annotations because our reference genomes were solely used to identify sequence variants (i.e., SNPs) for comparisons of regional populations of *P. citrophthora* and *P. syringae*.

Our analyses indicate that genetically distinct populations of both *P. citrophthora* and *P. syringae* are present in California. Subpopulations of both species, mostly defined by their geographic origin, were clearly identified by UPGMA, PCA, DAPC, fastStructure, and MSN analyses. FastStructure did not resolve the major cluster of *P. citrophthora* isolates but identified subpopulations in *P. syringae*. There was limited overlap between geographic subpopulations, indicating restricted airborne dispersal of the inoculum or possible anthropogenic movement of these soilborne *Phytophthora* species. Sampling regions were separated by distances of more than 150 km and by major urban regions or mountainous areas without agriculture. The limited overlap in clustering of geographic populations that was more pronounced in *P. syringae* could be explained by movement of harvested fruit to distant packinghouses that may be located in a different growing area. Decayed fruit, including those with brown rot, are sorted out in the packinghouse and are sometimes dumped in landfill areas or in orchards where they can become an inoculum source.

Overall, the population structure of *P. citrophthora* across three main citrus-growing regions of California can be described as consisting of closely related but distinct regional populations with relatively low intra-group diversity. This finding is consistent with the heterothallic nature of this species that requires the presence of two mating types. Additionally, oogonia, antheridia, and oospores have not been observed in nature, and it is assumed that the sexual stage is not a common part of the life cycle (Erwin and Ribeiro 1996). Oospores of *P. citrophthora* have been produced only in laboratory matings of certain isolates or in matings with other species (Erwin and Ribeiro 1996). Asexual reproduction in this species can explain the observed similarity among isolates and groups in our study. Still, a more diverse cluster of five CV isolates was identified, and these isolates possibly originated by hybridization, as has been described for numerous *Phytophthora* species (Van Poucke et al. 2021) or by introductions from different hosts or from other citrus-growing areas.

In contrast to the mostly closely related subpopulations of *P. citrophthora*, *P. syringae* exhibited much more variation with clonal lineages existing within subpopulations. Similar as for *P. citrophthora*, subpopulations of *P. syringae* were often associated with specific growing regions, and some MLGs were present in two or all three regions. A higher variation in *P. syringae* can be expected because sexual reproduction likely occurs readily. Higher variability may have been responsible for this species becoming prevalent in citrus during the last 40 years. *P. syringae* is homothallic but capable of outbreeding; still, this limits the possibility of generating new genotypes as compared with an outcrossing species if both mating types are present in a single population.

Both species reproduce mostly asexually by formation of sporangia with zoospores that are the main infective propagules. This asexual reproduction can explain the clonality (I_A values > 0) that was inferred in all sampled geographic populations of both species. The overall I_A for both *P. citrophthora* and *P. syringae* was found to be greater than their respective simulated clonal data sets. Subpopulations of both species exhibited a range of calculated I_A values. For *P. citrophthora*, I_A values for IE and VE populations were closest to that of the simulated semi-clonal data set, whereas the I_A value for the CV population was clearly above that of the simulated clonal data set. For *P. syringe*, the I_A value for the CV population was closest to that of the semi-clonal data set, whereas those of IE and noncitrus populations were closest to that of the simulated clonal dataset, and that of the VE population was above that of the simulated clonal data set. These results reflect our current understanding of the reproduction and lifestyle of the homothallic *P. syringae*, where both sexual and clonal reproduction are thought to readily occur in the environment. The MSN generated for *P. syringae* reflected the expected structure, with VE isolates belonging to the fewest MLGs and CV (and noncitrus) isolates being the most diverse and fragmented into the most unique MLGs. Still, sample size may have some impact on these results, as only 14 VE isolates were available for analyses as compared with 80 and 36 isolates for the CV and IE regions, respectively. *P. syringae* is more prevalent in colder environments and, therefore, is expected to be more common in the inland areas of California (i.e., the CV and IE regions) where most of the isolates were obtained.

In contrast, the comparatively low I_A values calculated for the IE and VE populations of *P. citrophthora* are more difficult to explain because *P. citrophthora* is not thought to reproduce sexually in the environment. Our UPGMA dendrogram depicting these populations (Fig. 3A) shows extremely little genetic distance among individuals belonging to the IE and VE populations. Furthermore, the MSN plot generated for *P. citrophthora* using a distance threshold produced MLGs that we believe are accurate depictions of the clonal structure within these populations. It is possible that factors not considered may impact this method of simulating I_A values for a data set. Additional investigations are needed to assess and characterize the mode of reproduction in natural populations of *P. citrophthora*.

An important finding of our study is that citrus isolates of *P. syringae* were clearly differentiated from those from noncitrus tree crop hosts. This is especially noteworthy for isolates from the CV, where growing areas of different tree crop hosts often overlap. In rainy winter seasons, *P. syringae* can be a serious almond pathogen causing trunk and crotch cankers (i.e., aerial Phytophthora disease; Browne and Viveros 1999), and there has been a concern that the extensive almond cultivation contributed to the emergence of *P. syringae* as a major citrus brown rot pathogen. This hypothesis, however, is not supported by our data with a relatively limited sampling of noncitrus isolates, and the source of the citrus *P. syringae* inoculum remains unknown. Noncitrus isolates were genetically diverse. They grouped on distinct branches of the UPGMA dendrogram and represented multiple MLGs in the MSNs. More sampling is warranted to describe the population structures of noncitrus isolates of *P. syringae*.

Several measures were calculated to characterize differentiation within species and regional populations. For *P. citrophthora*, a low overall F_{ST} value (Table 5), as well as low G_{ST}, G''_{ST}, and Jost's D

TABLE 6. Genetic distance statistics of *Phytophthora citrophthora* and *P. syringae*

Species	Comparison of populations	G_{ST}	G''_{ST}	Jost's D
P. citrophthora	Central Valley–Ventura Co.	0.0094	0.0202	0.0017
	Central Valley–Inland Empire	0.0072	0.0157	0.0013
	Ventura Co.–Inland Empire	–0.0045	–0.0094	–0.0005
P. syringae	Central Valley[a]–Ventura Co.	0.2397	0.4328	0.0752
	Central Valley[a]–Inland Empire	0.0732	0.1497	0.0155
	Ventura Co.–Inland Empire	0.2601	0.4602	0.0807
	Central Valley[a]–noncitrus	0.4295	0.7121	0.2787
	Inland Empire–noncitrus	0.4406	0.7221	0.2843
	Ventura Co.–noncitrus	0.4032	0.6932	0.2787

[a] Only isolates from citrus were included for the Central Valley population of *P. syringae*.

values (Table 6), suggests relatively weak differentiation between populations from the three sampling regions. For *P. syringae*, the relatively high overall F_{ST} (Table 5) indicates greater differentiation between populations. The G_{ST}, G''_{ST}, and Jost's D values (Table 6) were all higher than for *P. citrophthora*. Differentiation was lowest for CV and IE isolates, which reflects the fact that fruit from the IE is more commonly shipped to the large CV packinghouses, whereas fruit shipping between VE and the CV or between VE and the IE is less common.

The validity of some of the statistics used in our study has been criticized, as the field of population genetics/genomics is continuing to develop. Some authors (Gerlach et al. 2010; Hedrick 2005; Jost 2008; Meirmans and Hedrick 2011) argue that the original F_{ST}/G_{ST} statistics do not take population size into account and are dependent on within-population diversity. Subsequently, these authors proposed various alternative statistics, such as G'_{ST}, $G''_{ST'}$, and Jost's D, that were claimed to more accurately measure "true differentiation." Others, however, have found flaws in these newly proposed methods, claiming that on their own, they do not provide full insight into the factors affecting population differentiation and structure and that they should be viewed in the context of G_{ST}, not as replacements for it (Verity and Nichols 2014; Whitlock 2011). Yet another study (Alcala and Rosenberg 2019) assessed the constraints imposed on these metrics by allele frequencies and found them to often produce similar results, and in cases where differences were observed, those differences could be explained by features of the data set more greatly affecting one metric than the others. Although it is difficult to assess which of the metrics (G_{ST}, $G''_{ST'}$, or Jost's D) provides the "most accurate" measurement of "true differentiation," the results obtained using all metrics were consistent for the two species of *Phytophthora* studied and provide confidence in the accuracy of our results.

In summary, the distinct lifestyles of the two main species of *Phytophthora* currently causing brown rot of citrus in California are reflected by the population structures described in this study. Both lifestyles, that is, heterothallic but likely only asexually reproducing (*P. citrophthora* with overall high I_A values and an MSN with limited genetic distance among fewer MLGs that is indicative of clonality) or homothallic and sexually and asexually reproducing (*P. syringae* with overall high I_A values and MSN graphs with greater genetic distance and more unique single-isolate MLGs based on region and host that is indicative of some degree of clonality but with greater diversity), are successful strategies in the survival of each species in environmentally diverse citrus-production areas of the state. The genetic diversity of both species that we identified indicates that species evolution and development of regional subpopulations are ongoing and that they will continue to adapt to changing environments. Thus, as a quarantine pathogen in some export countries, *P. syringae* potentially will cause more damage to California citrus production in the future.

Acknowledgments

We thank G. Browne for providing several isolates of *Phytophthora* spp. from noncitrus hosts, W. Hao and R. Belisle for their assistance with DNA extractions, A. Shands for his guidance in genomic data analysis, and M. Brandenburg for his help in orchard sampling.

Literature Cited

Adaskaveg, J. E., and Förster, H. 2014. Integrated postharvest strategies for management of Phytophthora brown rot of citrus in the United States. Pages 123-131 in: Post-Harvest Pathology, Plant Pathology in the 21st Century. D. Prusky and M. L. Gullino, eds. Springer International Publishing, Cham, Switzerland.

Alcala, N., and Rosenberg, N. A. 2019. G'_{ST}, Jost's D, and F_{ST} are similarly constrained by allele frequencies: A mathematical, simulation, and empirical study. Mol. Ecol. 28:1624-1636.

Anonymous. 2022. California Agricultural Statistics Review. www.cdfa.ca.gov/Statistics/PDFs/2022_Ag_Stats_Review.pdf

Aronesty, E. 2013. Comparison of sequencing utility programs. Open Bioinform. J. 7:1-8.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., and Pevzner, P. A. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19:455-477.

Browne, G. T., and Viveros, M. A. 1999. Lethal cankers caused by *Phytophthora* spp. in almond scions: Specific etiology and potential inoculum sources. Plant Dis. 83:739-745.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., and Durbin, R. 2011. The variant call format and VCFtools. Bioinformatics 27:2156-2158.

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., and Li, H. 2021. Twelve years of SAMtools and BCFtools. Gigascience 10:giab008.

Erwin, D. C., and Ribeiro, O. K. 1996. *Phytophthora* Diseases Worldwide. American Phytopathological Society, St. Paul, MN.

FAO. 2021. Citrus fruit statistical compendium 2020. Rome. https://www.fao.org/3/cb6492en/cb6492en.pdf

Feld, S. J., Menge, J. A., and Pehrson, J. E. 1979. Brown rot of citrus: A review of the disease. Citrograph 64:101-106.

Flynn, J. M., Hubley, R., Goubert, C., Rosen, J., Clark, A. G., Feschotte, C., and Smit, A. F. 2020. RepeatModeler2 for automated genomic discovery of transposable element families. Proc. Natl. Acad. Sci. U.S.A. 117: 9451-9457.

Francis, R. M. 2017. pophelper: An R package and web app to analyze and visualize population structure. Mol. Ecol. Resour. 17:27-32.

Geisseler, D., and Horwath, W. R. 2016. Citrus production in California. https://apps1.cdfa.ca.gov/FertilizerResearch/docs/Citrus_Production_CA.pdf

Gerlach, G., Jueterbock, A., Kraemer, P., Deppermann, J., and Harmand, P. 2010. Calculations of population differentiation based on G_{ST} and D: Forget G_{ST} but not all of statistics! Mol. Ecol. 19:3845-3852.

Graham, J., and Feichtenberger, E. 2015. Citrus *Phytophthora* diseases: Management challenges and successes. J. Citrus Pathol. 2:1-11.

Graham, J. H., and Menge, J. A. 2000. Phytophthora-induced diseases. Pages 12-15 in: Compendium of Citrus Diseases. L. W. Timmer, S. M. Garnsey, and J. H. Graham, eds. American Phytopathological Society, St. Paul, MN.

Gruber, B., Unmack, P. J., Berry, O. F., and Georges, A. 2018. dartR: An R package to facilitate analysis of SNP data generated from reduced representation genome sequencing. Mol. Ecol. Resour. 18:691-699.

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. 2013. QUAST: Quality assessment tool for genome assemblies. Bioinformatics 29:1072-1075.

Hao, W., Miles, T. D., Martin, F. N., Browne, G. T., Förster, H., and Adaskaveg, J. E. 2018. Temporal occurrence and niche preferences of *Phytophthora* spp. causing brown rot of citrus in the Central Valley of California. Phytopathology 108:384-391.

Hedrick, P. W. 2005. A standardized genetic differentiation measure. Evolution 59:1633-1638.

Jombart, T. 2008. adegenet: A R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403-1405.

Jost, L. 2008. G_{ST} and its relatives do not measure differentiation. Mol. Ecol. 17:4015-4026.

Kamvar, Z. N., Tabima, J. F., and Grünwald, N. J. 2014. *Poppr*: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2:e281.

Klotz, L. J. 1978. Fungal, bacterial, and nonparasitic diseases and injuries originating in the seedbed, nursery, and orchard. Pages 1-66 in: The Citrus Industry. Vol. IV. W. Reuther, E. C. Calavan, and G. E. Carman, eds. University of California, Division of Agricultural and Sciences, Oakland, CA.

Li, H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv: 1303.3997.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078-2079.

Manni, M., Berkeley, M. R., Seppey, M., and Zdobnov, E. M. 2021. BUSCO: Assessing genomic data quality and beyond. Curr. Protoc. 1:e323.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M. A. 2010. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20:1297-1303.

Meirmans, P. G., and Hedrick, P. W. 2011. Assessing population structure: F_{ST} and related measures. Mol. Ecol. Resour. 11:5-18.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. U.S.A. 70:3321-3323.

Nei, M., and Chesser, R. K. 1983. Estimation of fixation indices and gene diversities. Ann. Hum. Genet. 47:253-259.

R Core Team. 2022. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Raj, A., Stephens, M., and Pritchard, J. K. 2014. fastSTRUCTURE: Variational inference of population structure in large SNP data sets. Genetics 197:573-589.

Ribeiro, O. K. 1978. A Source Book of the Genus *Phytophthora*. J. Cramer, Vaduz.

Shakya, S. K., Grünwald, N. J., Fieland, V. J., Knaus, B. J., Weiland, J. E., Maia, C., Drenth, A., Guest, D. I., Liew, E. C. Y., Crane, C., Chang, T.-T., Fu, C.-H., Minh Chi, N., Quang Thu, P., Scanu, B., von Stowasser, E. S., Durán, Á., Horta Jung, M., and Jung, T. 2021. Phylogeography of the wide-host range panglobal plant pathogen *Phytophthora cinnamomi*. Mol. Ecol. 30:5164-5178.

Tabima, J. F., Gonen, L., Gómez-Gallego, M., Panda, P., Grünwald, N. J., Hansen, E. M., McDougal, R., LeBoldus, J. M., and Williams, N. M. 2021. Molecular phylogenomics and population structure of *Phytophthora pluvialis*. Phytopathology 111:108-115.

Tarailo-Graovac, M., and Chen, N. 2009. Using RepeatMasker to identify repetitive elements in genomic sequences. Curr. Protoc. Bioinformatics 25:4.10.1-4.10.14.

Van Poucke, K., Haegeman, A., Goedefroit, T., Focquet, F., Leus, L., Jung, M. H., Nave, C., Redondo, M. A., Husson, C., Kostov, K., Lyubenova, A., Christova, P., Chandelier, A., Slavov, S. de Cock, A., Bonants, P., Werres, S., Palau, J. O., Marçais, B., Jung, T., Stenlid, J., Ruttink, T., and Heungens, K. 2021. Unravelling hybridization in *Phytophthora* using phylogenomics and genome size estimation. IMA Fungus 12:16.

Verity, R., and Nichols, R. A. 2014. What is genetic differentiation, and how should we measure it—G_{ST}, D, neither or both? Mol. Ecol. 23:4216-4225.

Vogel, G., Gore, M. A., and Smart, C. D. 2021. Genome-wide association study in New York *Phytophthora capsici* isolates reveals loci involved in mating type and mefenoxam sensitivity. Phytopathology 111:204-216.

Whitlock, M. C. 2011. G'_{ST} and D do not replace F_{ST}. Mol. Ecol. 20: 1083-1091.

Wickham, H. 2016. ggplot2: Elegant Graphics for Data Analysis (3e). Springer Verlag, New York.

Winter, D. J. 2012. MMOD: An R library for the calculation of population differentiation statistics. Mol. Ecol. Resour. 12:1158-1160.

Xu, S., Chen, M., Feng, T., Zhan, L., Zhou, L., and Yu, G. 2021. Use ggbreak to effectively utilize plotting space to deal with large datasets and outliers. Front. Genet. 12:774846.

Xu, S., Li, L., Luo, X., Chen, M., Tang, W., Zhan, L., Dai, Z., Lam, T. T., Guan, Y., and Yu, G. 2022. *Ggtree*: A serialized data object for visualization of a phylogenetic tree and annotation data. iMeta 1:e56.

Zhang, F., Chen, H., Zhang, X., Gao, C., Huang, J., Lü, L., Shen, D., Wang, L., Huang, C., Ye, W., Zheng, X., Wang, Y., Vossen, J. H., and Dong, S. 2021. Genome analysis of two newly emerged potato late blight isolates sheds light on pathogen adaptation and provides tools for disease management. Phytopathology 111:96-107.