**Population Biology** 



# Genetic Structure and Subclonal Variation of Extant and Recent U.S. Lineages of *Phytophthora infestans*

Amanda Saville and Jean Beagle Ristaino<sup>+</sup>

Department of Entomology and Plant Pathology, North Carolina State University, Raleigh, NC 27695 Accepted for publication 30 April 2019.

### ABSTRACT

The oomycete *Phytophthora infestans* is an important plant pathogen on potato and tomato crops. We examined the genetic structure of extant 20th and 21st century U.S. lineages of *P. infestans* and compared them with populations from South America and Mexico to examine genetic relationships and potential sources of lineages. US-23, currently the most prevalent lineage detected in the United States, shared genetic similarity primarily with the BR-1 lineage identified in the 1990s from Bolivia and Brazil. Lineages US-8, US-14, and US-24, predominantly virulent on potato, formed a cluster distinct from other U.S. lineages. Many of the other U.S. lineages shared significant genetic similarity with Mexican populations. The US-1 lineage, dominant in the mid-20th century,

Potato late blight is one of the most destructive diseases of solanaceous crops. *Phytophthora infestans* (Mont) de Bary caused the Irish potato famine of the 1840s and is a global threat on potato and tomato crops. Knowledge of the population structure of the pathogen and its relation to phenotypic characteristics, such as fungicide sensitivity or aggressiveness, is important to develop effective management strategies for the disease (Saville et al. 2015).

*P. infestans* is a heterothallic oomycete with both sexual and asexual reproductive cycles. With few exceptions (e.g., Toluca Valley, Mexico; Scandinavia; and The Netherlands [Brurberg et al. 2011; Drenth et al. 1993; Fry et al. 2015; Yuen and Andersson 2013]), the asexual reproductive cycle dominates, resulting in the development of distinct clonal lineages that can be characterized by features, such as aggressiveness, fungicide sensitivity, and host preference (Saville et al. 2015). In the United States, even when both mating types (A1 and A2) are present in a region, the establishment of a stable sexually reproducing population is rare (Danies et al. 2014; Hu et al. 2012). Multiple clonal lineages have been found in the United States since the 1990s, revealing a history of the displacement of lineages over time (Fry and Goodwin 1997; Hu et al. 2012).

The first disease caused by *P. infestans* was documented in the United States in 1843 and Europe in 1845 (Bourke 1964; Teschemacher 1845). Genetic analysis using simple sequence repeats (SSRs) of *P. infestans* from herbarium samples from 19th century historic outbreaks revealed the presence of a single dominant clonal lineage called FAM-1 (Herb-1 mitochondrial

<sup>†</sup>Corresponding author: J. B. Ristaino; Jean\_Ristaino@ncsu.edu

**Funding:** This work was supported by National Institute of Food and Agriculture, U.S. Department of Agriculture (https://nifa.usda.gov/) subaward number 2011-68004-30154 (to J. B. Ristaino).

\*The *e*-Xtra logo stands for "electronic extra" and indicates that seven supplementary figures and 11 supplementary tables are published online.

The author(s) declare no conflict of interest.

© 2019 The American Phytopathological Society

clustered with US-1 lineages from Peru. A survey of the presence of RXLR effector *PiAVR2* revealed that some lineages carried *PiAVR2*, its resistance-breaking variant *PiAVR2-like*, or both. Minimum spanning networks developed from simple sequence repeat genotype datasets from USABlight outbreaks clearly showed the expansion of US-23 over a 6-year time period and geographic substructuring of some lineages in the western United States. Many clonal lineages of *P. infestans* in the United States have come from introductions from Mexico, but the US-23 and US-1 lineages were most likely introduced from other sources.

Keywords: ecology, epidemiology

haplotype) in the 19th century that caused disease in both the United States and Europe (Saville et al. 2016). These data and others suggest migration of the pathogen between the two continents from a similar point of origin (Martin et al. 2013, 2014; Saville et al. 2016; Yoshida et al. 2013). The US-1 lineage (Ib mitochondrial haplotype) emerged in the 1930s in the United States, and the historic FAM-1 lineage subsequently declined (Saville et al. 2016). The US-1 clonal lineage remained the dominant lineage in the United States until the 1980s. Late blight outbreaks in the United States caused by the US-1 lineage were well managed with the fungicide mefenoxam until the late 1980s, when new lineages of the pathogen emerged that were insensitive to mefenoxam (Goodwin et al. 1996). In the 1980s, new genotypes of P. infestans, including US-6 (mating type A1), US-7 (mating type A2), US-8 (mating type A2), and US-11, emerged out of Mexico (Goodwin 1997; Goodwin et al. 1995a, 1998). The most common putative recombinant, US-11, was hypothesized to be the progeny of US-6 and US-7 lineages (Gavino et al. 2000). It was proposed that the majority of these new lineages emerged as the result of sexual recombination occurring in Mexico, whereas some arose as clonal derivatives of earlier lineages (Goodwin et al. 1998). Today, the US-11 lineage still occurs in fields in the Pacific Northwest and Florida.

The population structure of *P. infestans* underwent a major shift in 2009, when a new lineage, named US-22, was identified (Hu et al. 2012). A combination of ideal weather, the widespread distribution of infected tomato transplants through retail stores, and a lack of public familiarity with symptoms of late blight enabled the pathogen to spread, resulting in a major pandemic in the northeastern United States. Surprisingly, the US-22 lineage was sensitive to the fungicide mefenoxam, whereas the previous US-8 and US-11 lineages were insensitive to the fungicide (Saville et al. 2015). Two other lineages, US-23 and US-24, were also first identified in 2009, although at low levels, and they were also sensitive to mefenoxam (Hu et al. 2012). After the 2009 epidemic, the disease alert and forecasting system USABlight (www.usablight.org) funded by the U.S. Department of Agriculture National Institute of Food and Agriculture was launched to monitor genotype populations of P. infestans in the United States on a larger scale (Fry et al. 2013). Maps from some of the outbreaks in the United States in prior years and records of SSR genotypes and mefenoxam sensitivity of lineages of *P. infestans* have been published (Fry et al. 2015; Hansen et al. 2016; Saville et al. 2015). This information can be used to predict potential migration routes of the pathogen. Populations of *P. infestans* have undergone another population shift in the last 5 years. The older mefenoxam-resistant lineage, US-8 on potato, has declined, and it has been replaced by the mefenoxam-sensitive lineage US-23, possibly because of reduced mefenoxam usage or increased fitness of US-23 (Fall et al. 2015; Fry et al. 2015; Hu et al. 2012; Saville et al. 2015).

Effectors of P. infestans change in response to host R gene deployment (Vleeshouwers et al. 2011). The suite of effector haplotypes present in historic P. infestans was smaller and different from those of modern P. infestans, and most of the effectors that are important in virulence in modern lineages were absent in the historic lineages (Martin et al. 2013; Yoshida et al. 2013). Variations in haplotypes of specific effector genes among lineages of P. infestans, including Avr3a and PiAVR2, have been reported in Europe (Cárdenas et al. 2011; Gilroy et al. 2011) and may point to geographic sources of lineages: some R genes are derived from Mexico, whereas others are derived from South America (Montarry et al. 2010; Vleeshouwers et al. 2011).

Several studies have utilized next generation sequencing, genotyping by sequencing, and genome-wide analysis of new and older lineages of the pathogen (Hansen et al. 2016; Knaus et al. 2016; Martin et al. 2013, 2016). We sequenced >70 genomes of P. infestans from newer and historic outbreaks and documented the ancestry of the famine era lineage and evolution of the US-1 lineage (Martin et al. 2016). The modern US-8, US-22, and US-23 lineages were placed in a derived clade than US-1 (Martin et al. 2013). Genotyping by sequencing was used to examine the genetic relationships of four modern lineages and document subclonal variation (US-8, US-11, US-23, and US-24) (Hansen et al. 2016). A close evolutionary relationship between US-11 and US-18, a wellsupported clade for US-1, and migration from Mexico were reported in another study (Knaus et al. 2016), but relationships among other extant lineages were not discerned. The genetic diversity of the pathogen within Mexican populations of P. infestans has been examined (Wang et al. 2017). Those workers suggested that all U.S. lineages came from Mexico and that different lineages shared genetic similarity within different regions of Mexico. However, that study did not include all extant U.S. lineages or lineages from other sources, including South American sources. A comprehensive study of the genetic relationships of all extant 20th and 21st century U.S. genotypes of P. infestans, including rare genotypes detected in the mid-1990s and widespread lineages, has not been done. The primary objectives of our work were to (i) examine the genetic structure among all extant and emerged U.S. lineages of P. infestans, (ii) test for the presence of recombination within these U.S. lineages, and (iii) infer potential migration pathways of the lineages within and into the United States from Mexico and/or South American countries.

# MATERIALS AND METHODS

Sampling. A total of 190 P. infestans isolates were studied using both multilocus sequencing and microsatellite genotyping and included extant isolates from 18 of the 24 known U.S. lineages from the United States and isolates from Mexico and South America (Bolivia, Brazil, Ecuador, and Peru). The dates (1984 to 2014) and locations where the isolates were obtained are indicated in Supplementary Table S1.

Because some rare lineages were found in very limited numbers or locations and sometimes only in a single field (Goodwin et al. 1998), those lineages are represented by one or two samples in our study (US-12, US-14, US-15, US-16, US-17, and US-19). Restriction fragment length polymorphisms (RFLPs) fingerprints for all available U.S. lineages as well as South American clonal lineages EC-1, PE-3, and PE-7 were obtained from previously published sources (Forbes et al. 1998; Garry et al. 2005; Goodwin et al. 1998; Hu et al. 2012; Schultz et al. 2010; Wangsomboondee et al. 2002) (Table 1). We also used samples collected from outbreaks (n = 755)archived in the USABlight system from 2011 to 2016 (Fig. 1) to examine the genetic structure of recently emerged U.S. populations of P. infestans.

TABLE 1. Summary of previously reported phenotypes and genotypes of lineages of Phytophthora infestans present in the United States

				Allozyme genotype						
Lineage	First U.S. report	Hosta	Mating type	Gpi	Pep	mtDNA	Mefenoxam sensitivity <sup>b</sup>	RG57 RFLP <sup>c</sup>		
US-1	1931 <sup>d</sup>	P/T	A1	86/100	92/100	Ib	S	1,3,4,5,7,9,10,13,14,16,20,21,24,25		
US-6	1979 <sup>e</sup>	P/T	A1	100/100	92/100	IIb	S/R	1,3,4,5,6,7,10,13,14,18,20,21,24,25		
US-7	1992 <sup>f</sup>	P/T	A2	100/111	100/100	Ia	S/I/R	1,4,5,10,13,14,16,18,20,21,24,25		
US-8	1992 <sup>f</sup>	P/T	A2	100/111/122	100/100	Ia	I/R	1,5,10,13,14,16,20,21,23,24,25		
US-11	1994 <sup>e</sup>	P/T	A1	100/100/111	100/100	IIb	R	1,3,5,6,7,10,13,14,16,18,20,21,24,25		
US-12	1994 <sup>e</sup>	Т	A1	100/111	92/100	IIb	R	1,5,10,13,14,18,20,21,24,25		
US-13	1994 <sup>e</sup>	Т	A2	100/100	100/100	ND <sup>g</sup>	R	-		
US-14	1994 <sup>e</sup>	Р	A2	100/122	100/100	Ia	R	1,5,10,13,14,16,20,21,23,24,25		
US-15	1994 <sup>e</sup>	Т	A2	100/100	92/100	Ia	S	1,5,10,13,14,15,18,20,21,24,25		
US-16	1994 <sup>e</sup>	Т	A1	100/111	100/100	Ia	R	1,5,6,10,13,14,16,18,20,21,24,25		
US-17	1996 <sup>e</sup>	P/T	A1	100/122	100/100	ND <sup>g</sup>	R	1,3,7,13,14,16,18,20,21,24,25		
US-18	1994 <sup>h</sup>	Т	A2	100/100	92/100	Ia	S/I	1,5,10,13,14,16,20,21,24,25		
US-19	1997 <sup>h</sup>	Т	A2	100/100	92/100	Ia	S/I	1,3,5,7,13,14,16,20,21,24,25		
US-20	2005 <sup>i,j</sup>	Т	A2	100/100	100/100	Ia	S/I/R	1,3,5,7,10,13,14,16,18,20,21,24,25		
US-21	2006 <sup>i,j</sup>	Т	A2	100/122	100/100	Ia	S/I/R	1,5,10,13,14,18,20,21,24,25		
US-22	2007 <sup>j</sup>	P/T	A2	100/122	100/100	Ia	S/I	1,5,13,14,16,20,21,24,25		
US-23	2009 <sup>j</sup>	P/T	A1	100/100	100/100	Ia	S/I <sup>k</sup>	1,2,5,6,10,13,14,17,20,21,24,24a,25		
US-24	2009 <sup>j</sup>	P/T	A1	100/100/111	100/100	Ia	S/I	1,3,5,7,10,13,14,16,20,21,23,24,25		

<sup>a</sup> P, potato; T, tomato.

<sup>b</sup> I, mefenoxam intermediate; R, mefenoxam resistant; S, mefenoxam sensitive.

<sup>c</sup> Restriction fragment length polymorphisms (RFLPs) given here. Raw data from simple sequence repeat (SSR) are archived on USABlight.org.

<sup>d</sup> Report from Saville et al. (2016).

Report from Goodwin et al. (1995).

f Report from Goodwin et al. (1998).

<sup>g</sup> The mtDNA (mitochondrial DNA) haplotype for US-13 and US-17 is reported as Ia. ND, not determined.

<sup>h</sup> Report from Wangsomboondee et al. (2002). i

Report from Schultz et al. (2010).

Report from Hu et al. (2012).

<sup>k</sup> Isolates of US-23 intermediate in sensitivity to metalaxyl reported in Matson et al. (2015).

**DNA extraction, PCR, and sequencing.** DNA was extracted using either a hexadecyltrimethylammonium bromide method (May and Ristaino 2004) or a modification of the quick sodium hydroxide extraction method described by Wang et al. (1993). The sodium hydroxide extraction was done by grinding tissue in 90  $\mu$ l of 0.5 N sodium hydroxide and immediately transferring 3  $\mu$ l of the ground sample to 300  $\mu$ l of 100 mM Tris, pH 8.0.

Three nuclear loci (*Piypt1*, *PiAVR2*, and *PiAVR2-like*) were sequenced. The PCR primers, their location on the target sequence, and the source are shown in Supplementary Table S2. For the *Piypt1* gene, two regions were amplified, including intron 1 (223 bp with IRF/IRR) located in the 5' untranslated region of the gene and a 600-bp portion (with RASF/RASR) covering part of exon 3, exon 4, exon 5, part of exon 6, and introns 3 and 4 (Gómez-Alpizar et al.



Fig. 1. Outbreaks of late blight as collected through USABlight from 2011 to 2016. A, Map of locations of recorded outbreaks of late blight plotted by lineage collected through USABlight from 2011 to 2016. Disease occurrence is color coded by genotype detected. Larger points are indicative of multiple records from a location. The map was constructed using Tableau Desktop (Tableau Software). Frequencies of genotypes detected from outbreak data obtained through USABlight from 2009 to 2016 from **B**, potato and **C**, tomato hosts.

2007). The avirulence (AVR) gene *PiAVR2* and its virulent variant *PiAVR2-like* were amplified using primers AVR2F1/AVR2R2 and AVR2F6/AVR2R6, respectively (Gilroy et al. 2011). These sets of primers amplify neutral regions of two loci associated with avirulent and virulent phenotypes in *P. infestans* to the R2 resistance gene in potato (Gilroy et al. 2011).

Two 50-µl master mixes were prepared for each sample. Each reaction contained 5 µl of 10× PCR buffer (Genesee), 2.5 µl dNTP buffer (2 mM per nucleotide), 2 µl each 10 µM forward and reverse primer, 1.8 µl of MgCl<sub>2</sub> (50 mg/ml), 0.25 µl of BSA (20 mg/ml), 0.2 µl of Taq (5 U/µl; Genesee), and 5 to 10 ng of genomic DNA. Thermal cycling conditions for nuclear genes were 96°C (1 min); 35 cycles of 96°C (1 min), 55°C (1 min), and 72°C (2 min); and a final extension of 72°C (10 min). Products were purified with ExoSap (Affymetrix). Purified fragments were sequenced directly on both strands by using the same primers as those used in the initial amplification. Sequencing reactions were prepared using the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit and analyzed on an ABI PRISM 3730xl automated sequencer (Applied Biosystems).

**SSR genotyping.** *P. infestans* was genotyped using a modified version of the protocol for 12-plex single-sequence repeat genotyping (Danies et al. 2013; Li et al. 2013). The Qiagen Type-It Microsatellite PCR kit (Qiagen Corporation) was used for PCR reactions, and sample volumes were modified to run a 12.5- $\mu$ l reaction (Saville et al. 2016). Thermal cycling conditions followed the work by Danies et al. (2014). For fragment analysis, 1 to 2  $\mu$ l of PCR product was added to a 10.3- $\mu$ l reaction mix consisting of 10  $\mu$ l highly deionized formamide and 0.3  $\mu$ l of LIZ500 size standard (Applied Biosystems). Fragments were analyzed on an Applied Biosystems 3730xl DNA analyzer at the Genomic Sciences Laboratory at North Carolina State University. Alleles were scored manually using Peak Scanner 2 (Applied Biosystems) and rounded to the nearest whole number for analysis.

**Data analysis.** For specific analyses, samples were divided by geographic region, including the United States, Mexico, pooled South America, or country of origin in South America: Bolivia, Brazil, Ecuador, or Peru.

SSR genotyping analyses. Analysis of SSR genotypes was done using the standalone version of the program Structure v.2.3.4 (Pritchard et al. 2000). Before analysis, the data were clone corrected (identical genotypes reduced to one representative per population) using the R library poppr (Kamvar et al. 2014). The data were run using a 20,000 repeat burn in and 1,000,000 MCMC repeats. Because the data included representatives of a sexually reproducing population from Mexico, the admixture model was selected for analysis of the entire dataset. Datasets were run from K values 1 to 10 for 20 repetitions of each K value. The optimal K was predicted using the Evanno method in the web tool Structure Harvester (Earl and vonHoldt 2012). In addition, the optimal K was inferred through direct observation of cluster probabilities. All runs for the optimal K as well as surrounding K values were averaged using CLUMPP v. 1.1.2 (Jakobsson and Rosenberg 2007) and visualized with the program Distruct v. 1.1 (Rosenberg 2004). SSR genotypes were also visually examined for the presence of unique and rare alleles in a modification of the visual parentage exclusion test described by Danies et al. (2014). To further visualize groupings, a discriminant analysis of principal components (DAPC) was conducted using the R library Adegenet (Jombart 2008). An estimate of the number of principal components to utilize was generated using cross-validation tests. Rooted neighbor joining trees based on Bruvo's distance with a combination model (genome addition and gene loss) were generated using the R libraries poppr and pegas with 1,000 bootstrap iterations (Kamvar et al. 2014; Paradis 2010). To further examine distances between multilocus genotypes (MLGs), a minimum spanning network (MSN) was generated for all genotypes. Poppr was also used to infer population statistics, including the following indices: Shannon and Weiner index (Shannon 2001), Stoddart and Taylor's index (Stoddart and Taylor 1988), Simpson's index (Simpson 1949), Nei's unbiased gene diversity (Nei 1978), evenness (Grünwald et al. 2003; Ludwig 1988; Pielou 1975), and the index of association (Brown et al. 1980; Smith et al. 1993). To compensate for the changes in sample size that can affect Simpson's index ( $\lambda$ ), we used a size correction by multiplying  $\lambda$  by *N*/(*N*-1), where *N* is the sample size as described by Wang et al. (2017). MLGs and population statistics were generated for the dataset containing all U.S. lineages, South American and Mexican populations, and the outbreak dataset from USABlight. With the exception of population statistics, all analyses conducted in R utilized a complete dataset. To account for the presence of both diploid and triploid alleles, all data were adjusted to be triploid using methods in the poppr R library.

Gene sequence analysis. All statistical analyses of the nucleotide sequences were performed in SNAP Workbench version 2.0 (Monacell and Carbone 2014). All sequences were aligned manually and edited using BioEdit (Hall 1999). Multiple sequence alignment was also performed in Clustal X (Thompson et al. 1994). All polymorphisms were checked from the chromatograms, and heterozygous sites were determined. The programs PHASE (Stephens and Donnelly 2003) and SeqPHASE (Flot 2010) were utilized to predict genotypes for individual sequences of PiAVR2like. PHASE was run five times with different random seeds to ensure convergence of results. Sequences were collapsed into unique haplotypes using SNAP Map (Aylor et al. 2006) after removing insertions and deletions (indels) from each of the aligned multilocus datasets and excluding infinite sites violations. Base substitutions were categorized as phylogenetically informative or uninformative, transitions or transversions, and nonsynonymous (replacement) or synonymous amino acid changes in the coding region of each alignment. Resultant haplotype datasets were used to examine the overall support or conflict among the variable sites in the DNA sequence alignment. A site compatibility matrix was generated from each haplotype dataset using SNAP Clade (Aylor et al. 2006). Compatibility matrices were used to examine compatibility/incompatibility among all variable sites, with any resultant incompatible sites removed from the dataset for use in analyses where data are assumed to be fully compatible. The largest nonrecombining block for each locus was determined using cladeEx within SNAP Workbench (Bowden et al. 2008). Datasets were also evaluated using Kwarg (Lyngsø et al. 2005) for estimating the minimum number of recombination events and mapped into ancestral recombination graphs (ARGs). Conflicting data partitions or putative recombinant haplotypes were also excluded from additional analyses. Nonrecombining datasets were collapsed into unique haplotypes excluding infinite sites violations using SNAP Map. Three nonrecombining datasets were defined as follows: Pivpt1, PiAVR2, and PiAVR2-like.

**Neutrality tests.** After the largest nonrecombining block was defined for each locus, the DNA sequences were analyzed using Arlequin (Excoffier and Lischer 2010). For each locus in each population, the population mutation parameter per nucleotide site  $\theta$  using Watterson's estimate  $\theta_w$  (Watterson 1975) based on the number of segregating sites *s* and the average pairwise nucleotide diversity  $\pi$  (Tajima 1983) was estimated, and different tests of neutrality (Fu 1997; Tajima 1989) were performed to determine whether the data were consistent with the expectations of the neutral model of molecular evolution.

Gene flow with Hudson's statistics. Populations of *P. infestans* from South America, Mexico, and the US-23 lineage were tested for subdivision using Hudson's test statistics and the *Ypt1* gene. Our previous analysis of SSR data indicated clustering of most of the other U.S. lineages except US-1 with Mexican populations. The full *Ypt1* gene sequence, including instances of recombination, were utilized, because recombination events increase the power of these tests (Hudson 2000; Hudson et al. 1992a, b). The nearest neighbor statistic (*Snn*) was calculated for two separate comparisons of the

populations. In one, all of South America was regarded as a single population. In the other, populations from South America were divided by country (Bolivia, Brazil, Ecuador, and Peru), and data from each country were tested individually.

Analysis of potential migration scenarios. Migration patterns of the currently prevalent lineage US-23 were evaluated using approximate Bayesian comparison (ABC) as implemented in the program DIYABC version 2.0.4 (Cornuet et al. 2010). The tested scenarios evaluating potential migration paths for US-23 included direct divergence from South America or Mexico, admixture with South America and Mexico populations, and admixture between South America or Mexico and an unsampled population. Parameter range priors were based on previously published data and modified to accommodate US-23 (Saville et al. 2016) (Supplementary Table S3). Scenario probabilities were determined through comparison of the observed dataset to simulated datasets generated by DIYABC. A logistic regression of these differences was computed using 10 proportions of the simulated dataset as the dependent variable and corresponding differences between the observed and simulated datasets as the independent variable. The highest value calculated was taken as the scenario's overall probability. Confidence in the three highest scenarios was evaluated using type I and type II error tests, in which the data were compared against 500 simulated datasets and the number of times that the scenario in question was correctly or incorrectly applied to the data was determined.

# RESULTS

**Population structure.** A total of 110 MLGs were detected within the U.S. lineages and South American and Mexican populations (Supplementary Table S4). The greatest number of MLGs was observed among South America populations. The index of association ( $I_a$ ), a measurement of linkage disequilibrium, was calculated (Smith et al. 1993). South American populations had an  $I_a$ , furthest away from 0, indicating the greatest level of clonality ( $I_a = 2.79$ ). In contrast, Mexican populations had an  $I_a$  close to 0, indicating the greatest level of recombination ( $I_a = -0.065$ ). Under the standardized index of association ( $\bar{r}_d$ ), which incorporates the number of loci into the calculations, South American populations were farther from 0 ( $\bar{r}_d = 0.269$ ) than Mexican populations ( $\bar{r}_d = -0.006$ ).

A total of 123 SSR alleles were detected across all 12 microsatellite loci (average: 10.25 alleles per locus). The highest numbers of SSR alleles were found at PiG11 (n = 21), D13 (n = 24), and PinfSSR4 (n = 16). The SSR loci Pi04 and PinfSSR2 had the fewest number of alleles (n = 4) (Supplementary Table S5). Locus Pi63 had the lowest level of diversity based on the corrected Simpson's index (0.366) and shared the lowest level of expected heterozygosity with PinfSSR2 (0.32), whereas locus D13 had the highest level of diversity based on both parameters (0.95 for corrected Simpson's index and 0.91 for expected heterozygosity).

Visual examination of microsatellite calls revealed the presence of lineage-specific alleles. US-1 had the greatest number of loci with alleles unique to the lineage found at PiG11 (200), D13 (136), Pi70 (187 of 188), Pi63 (274 of 275), and PinfSSR2 (177). Triploid loci were observed in US-1, US-7, US-8, US-11, US-14, and US-23 as well as lineages from Ecuador and Peru.

We inferred the population structure based on SSR genotypes using both Structure Harvester and direct observation of probabilities. The optimal *K* value determined by Structure Harvester was K = 7, and the optimal *K* value determined by observation of probabilities was K = 6. At K = 6, the US-1 lineage was distinct and clustered with US-1 lineages from Peru (Fig. 2). US-8, US-14, US-20, and US-24 formed a cluster at K = 7. US-7, US-11, US12, US13, and US-15-US19 formed another group that was similar to each other and grouped with lineages from Mexico. US-21 and US-22, both tomato-specific lineages, formed another cluster at K = 7. The US-23 lineage clustered with South American lineage BR-1 from Bolivia and Brazil. Other South American isolates from Peru and Ecuador clustered together and were predominately the EC-1 genotype. The PE-3 genotype from Peru clustered with isolates from Bolivia and Brazil.

Brazil, and these lineages were distinct from other South American lineages (EC-1 and PE-3). The EC-1 lineages from Ecuador and Peru formed a clade, but differences among the EC-1 lineages from Ecuador were observed, suggesting subclonal variation (Fig. 3). In addition, potato lineages US-8, US-14, and US-24 formed a well-defined clade that was sister to Mexican lineages. US-1 formed a clade with Peruvian isolates also identified as US-1. Other lineages

SSR data were also plotted into a neighbor joining tree using the Bruvo's distance algorithm. Similar to the Structure analysis, US-23 isolates formed a sister clade with BR-1 lineages from Bolivia and



Fig. 2. Structure analyses of 12-plex simple sequence repeat loci from populations of *Phytophthora infestans*. All graphs were generated from a single run of Structure (*K* values 1 to 10 tested; burn-in chain of 20,000 repeats; Markov chain Monte Carlo run of 1,000,000 repeats; each *K* value repeated 20 times) and compiled using CLUMPP and Distruct.

that formed independent clades were US-18, US-20, US-21, and US-22. US-21 and US-22 group together as sister lineages.

Analysis of the SSR data based on eight principal components indicated the presence of multiple exclusive groups (Fig. 4). US-23 was distant from all other U.S. lineages but clustered near the Bolivian and Brazilian BR-1 lineage. Similarly, US-1 clustered closest to South American US-1 populations from Peru/Ecuador and not Mexican populations. The remaining U.S. lineages clustered with Mexican and not South American populations.

Examination of the lineages using an MSN revealed population structures similar to those seen in our other analyses. The center of the network graph consists of South American isolates from Ecuador and Peru genotyped either as EC-1 or PE-6 surrounded by three major branches (Fig. 5). One branch included all U.S. lineages except for US-1 and US-23 and included multiple Mexican MLGs. Within this branch, lineages US-8, US-14, and US-24 formed a distinct group with four Mexican MLGs. The US-6, US-7, and US-13 lineages also formed separate branches with Mexican lineages. A second branch included US-1, nine Peruvian MLGs, and one Mexican MLG, which was shared with the third branch. The nine Peruvian MLGs were contained within two subbranches, one consisting of MLGs of US-1 isolates from Peru and the other containing isolates identified as PE-3. The third major branch included US-23, Bolivian and Brazilian MLGs, and five Mexican MLGs (including one shared with the second branch).

Ten different migration scenarios were examined using SSR allele data and ABC analysis among the US-23, South America, and Mexico populations. The scenario with the highest probability (Scenario 1, P = 0.642) was chosen as the most likely model. This scenario was a model in which Mexican populations diverged first from a common ancestor that then diverged into US-23 and South America populations (Fig. 6A). Confidence in the scenario choice was evaluated by using simulated datasets to calculate error percentages between the three scenarios with the highest probabilities. Estimation of type I error revealed that 51.6% of simulated datasets using this scenario 2 when compared with the two scenarios with the next highest probabilities (Scenarios 9 and 10; type I error, 0.484) (Supplementary Fig. S1 and Supplementary Table S6).

**Population structure of U.S. lineages from USABlight, 2011 to 2016.** A total of 325 MLGs were detected within USABlight outbreak samples collected between 2011 and 2012, and 430 MLGs



Fig. 3. Neighbor joining tree of 12 simple sequence repeat loci using 18 U.S. lineages of *Phytophthora infestans* and samples from South America and Mexico. Distances were calculated using Bruvo's distance.

were detected from samples collected between 2013 and 2016. The greatest number of MLGs was observed within populations of US-23 during both time periods. This lineage also had the highest diversity indices between 2013 and 2016 (Shannon and Weiner, Stoddard and Taylor's, and corrected Simpson's index). The total number of clonal lineages decreased over time, and US-23 lineage became dominant on both potato and tomato over time (Fig. 1B and C).

Two MSNs were constructed for samples collected from 2011 to 2016 and genotyped by SSR from the USABlight archive (Fig. 7). Because samples collected between 2011 and 2012 were genotyped using different primers and fewer numbers of loci (Lees et al. 2006) than the 12-plex panel that was used between 2013 and 2016 (Li

et al. 2013), the data were analyzed separately. During the 2011 to 2012 period, multiple rare variants were observed as well as several lineages documented during the 2009 epidemic, including US-22, US-23, and US-24. US-8, US-24, and a variant of US-8 clustered together to form a distinct branch of the MSN (Fig. 7A). Haplotypes identified as US-22 cluster together, and they formed a distinct branch with two rare variants (new type C 2011 and new type G 2011). Haplotypes of US-11 also formed a distinct branch. US-23 began to emerge between 2011 and 2012 (Fig. 7A), and clearly, the number of haplotypes of US-23 expanded greatly with time (Fig. 7B). The occurrence of the US-23 lineage increased between 2013 and 2016 to become the dominant lineage present in the United



Fig. 4. Discriminant analysis of principal components plot of 12 simple sequence repeat loci from *Phytophthora infestans*. Lineages/populations of *P. infestans* are represented by color and symbol as well as inertia ellipses. The graph was generated using eight principal components. BOL, Bolivia; BR, Brazil; ECU, Ecudaor; MEX, Mexico; PER, Peru.

States during that time period (Fig. 1B and C). An increase in the number of different US-23 MLGs was observed mostly in the eastern United States in many states (Fig. 1A). US-8 and US-24 lineages formed distinct branches in the MSN (Fig. 7B) and were geographically substructured on the West Coast or in North Dakota, respectively (Fig. 1A). US-11 formed a separate branch of MLGs linked to US-7 lineages (Fig. 7B), and these lineages were mostly identified in the western United States and Florida (Fig. 1A). A few rare variants were identified that were linked to US-7, US-8, and US-24 lineages (Fig. 7A). The decline in the US-22 lineage and

increase in dominance of the US-23 MLG occurred over time (Supplementary Fig. S2). Shared MLGs were observed between tomato and potato, but more diverse MLGs were observed on tomato (Table 2) (Supplementary Fig. S3). Many diverse MLGs were observed in New York, where an ephemeral sexual population was described (Danies et al. 2014). Unique branches were observed in the MSN comprising lineages from the West Coast, including Oregon, Washington, and California (Supplementary Fig. S4).

**Nuclear sequence variability.** A total of 1,089 nucleotides were sequenced that corresponded to 766 nucleotides in *Piypt1* and



Fig. 5. Minimum spanning network of 12 simple sequence repeat loci from *Phytophthora infestans*. Node size is proportional to the number of isolates of *P. infestans* with the given multilocus genotype. Increasing thickness and color of the lines is proportional to the distance between two nodes. Only isolates for which all 12 loci could be obtained are included.

323 nucleotides in *PiAVR2* and *PiAVR2-like* (Supplementary Table S7). Allele sequences were used for the analysis because of the heterozygous sites in *Piypt1*, *PiAVR2*, and *PiAVR2-like*. There were two haplotype sequences for each isolate. In the case of a homozygous sequence, two identical sequences were used for the analysis. Because no recombining sites were identified in *PiAVR2* or *PiAVR2-like*, diversity statistics for these loci were calculated using the entire sequence. For *Piypt*, the segregating sites from the largest nonrecombining block were used.

Thirteen segregating nucleotide sites were identified in the *Piypt1* gene, resulting in a total of 14 haplotypes identified (Supplementary Table S8). The greatest number of haplotypes was found in the South America and Mexico populations (Supplementary Table S9). Nucleotide diversity ( $\pi$ ), the average number of nucleotide differences

per site between two sequences, was 0.378 for the pooled sample. The Watterson's estimate ( $\theta_w$ ) of population mean mutation rate for the pooled sample was 1.114. Lineages US-13 and US-14 had the highest nucleotide diversity ( $\pi$ ) and mean mutation rates ( $\theta_w$ ). All populations were determined to be neutral based on Tajima's *D* and Fu's *Fs* statistics.

One segregating nucleotide site was identified in the *PiAVR2* gene, and three were identified in the *PiAVR2-like* variant. Nucleotide diversity ( $\pi$ ) values for the pooled samples were 0.001 and 0.003 for *PiAVR2* and *PiAVR2-like*, respectively. The Watterson's estimate ( $\theta_w$ ) values of population mean mutation rate for the pooled samples were 0.180 and 0.418 for *PiAVR2* and *PiAVR2-like*, respectively. All populations were determined to be neutral based on Tajima's *D* and Fu's *Fs* statistics.



Fig. 6. Scenarios A, 1, B, 9, and C, 10 for the evolution of *Phytophthora infestans* from sampled populations using approximate Bayesian comparison. Posterior probabilities of the top three migration scenarios comparing populations from the US-23 lineage and the South American (SA) and Mexican (MEX) populations. Probabilities are the highest value of a logistic regression of data.



**Fig. 7.** Minimum spanning network of simple sequence repeat loci collected from genotyping of *Phytophthora infestans* samples submitted through the USABlight website from 2011 to 2016. **A**, Genotyping from 2011 to 2012 was conducted using the protocols of Lees et al. (2006) (n = 325), and **B**, genotyping from 2013 to 2016 was conducted using the protocols of Li et al. (2013) (n = 430). Node size is proportional to the number of isolates of *P. infestans* with the given multilocus genotype. Increasing thickness and color of the lines is proportional to the distance between two nodes.

Amplification of DNA from isolates using primers for *PiAVR2* and *PiAVR2-like* revealed the presence or absence of the avirulent *PiAVR2* or virulent *PiAVR2-like* variant corresponding to lineage (Fig. 8). US-1, US-6, US-11, US-12, US-15, US-16, US-19, US-21, and US-22 lineages were positive for the avirulent form of *PiAVR2*, whereas US-8, US-14, and US-20 were positive for the virulent form *PiAVR2-like*. Several lineages (US-7, US-13, US-17, US-18, US-23, and US-24) were positive for both the virulent and avirulent forms of the effector.

**Recombination among U.S. lineages.** The ARG for *Piypt1* showed five haplotypes (H2, H3, H4, H5, and H14) that were not generated through a recombination event (Supplementary Fig. S5). Most U.S. lineages were included within these haplotypes, except for US-1, US-12, US-15, US-16, US-17, US-20, and US-21. In addition, the nonrecombining haplotypes did not include any isolates from Mexico. All other haplotypes were generated through at least one ancestral recombination event. Recombination events led to the generation of haplotypes H7, H10, and H12, which included U.S. lineages US-11, US-15, US-16, US-18, and US-19 as well as the Mexican lineages.

No recombination was detected within *PiAVR2* and *PiAVR2-like* loci for these effectors (Supplementary Figs. S6 and S7).

**Gene flow.** When populations of *P. infestans* from South America were tested using Hudson's statistics as one population, subdivision was not detected between US-23 populations of *P. infestans* and populations in South America and Mexico (Supplementary Table S10). However, when populations of *P. infestans* from South America were analyzed by country, US-23 was not subdivided from *P. infestans* populations in Bolivia and Brazil, but it was subdivided from populations of *P. infestans* were subdivided from populations of *P. infestans* from Bolivia and Brazil but not those of Peru and Ecuador.

## DISCUSSION

We examined the genetic variation among all extant U.S. lineages of *P. infestans* to better understand the genetic structure of *P. infestans* in the United States over the past 40 years. The US-1 lineage is a particularly interesting case because of the longevity of its presence in the United States. The earliest known record of US-1 in the United States is from a herbarium specimen collected in 1931 in Texas (Saville et al. 2016). Although US-1 was a dominant lineage in the United States for 60 years, it declined in the mid-1990s, most likely because of its sensitivity to the fungicide mefenoxam. By comparison, all other U.S. lineages, with the exception of US-6, were detected in the 1990s in the United States, and many were resistant to mefenoxam (Table 1). The US-23 lineage is sensitive to the compound, but some shift to intermediate sensitivity has been observed (Matson et al. 2015).

The majority of lineages of *P. infestans* were isolated from tomato or on both potato and tomato. Only one lineage, US-14, has been observed exclusively on potato, whereas US-21 is a tomato-specific lineage (Hu et al. 2012). Other studies have suggested that host specificity in potato results from the gain or loss of avirulence genes (Gilroy et al. 2011; Oyarzun et al. 1998; Vleeshouwers et al. 2011).



**Fig. 8.** Detection of *PiAVR2* variants in U.S. lineages of *Phytophthora infestans*. Lineages where both *PiAVR2* and *PiAVR2-like* were detected are placed in the center region. P, U.S. genotype reported to infect potato; T, U.S. genotype reported to infect tomato.

TABLE 2. Genetic diversity statistics for populations of *Phytophthora infestans* based on microsatellite loci sorted by host, year, and region (data were collected from the USABlight database between 2011 and 2016)<sup>a</sup>

Population	Ν	MLG	eMLG (SE)	Н	G	Corrected $\lambda^{b}$	Evenness	Hexp	Ia	$\bar{r}_d$
Host										
2011-2012										
Potato	93	71	71 (0.00)	4.08	43.0	0.988	0.720	0.624	2.06	0.214
Tomato	232	143	66.1 (3.38)	4.40	25.4	0.965	0.304	0.596	2.15	0.226
All	325	199	69.4 (3.56)	4.75	36.0	0.975	0.304	0.608	2.04	0.212
2013-2016										
Potato	180	108	108 (0.00)	4.45	65.6	0.991	0.762	0.527	3.03	0.294
Tomato	250	137	107 (3.41)	4.55	59.3	0.987	0.625	0.509	2.74	0.270
All	430	207	112 (4.69)	4.88	76.7	0.989	0.581	0.517	2.88	0.280
State/province										
2011-2012										
New York	87	77	9.88 (0.338)	4.307	70.74	0.997	0.953	0.601	1.755	0.189
Florida	80	8	3.63 (0.907)	1.27	2.52	0.611	0.594	0.411	4.572	0.589
Pennsylvania	41	39	9.89 (0.319)	3.646	37.36	0.997	0.974	0.550	1.255	0.138
Maine	34	33	9.92 (0.272)	3.486	32.11	0.998	0.983	0.586	1.371	0.161
North Carolina	19	18	9.74 (0.440)	2.871	17.19	0.994	0.972	0.618	1.953	0.206
2013-2016										
New York	144	71	9.11 (0.873)	3.913	34.68	0.978	0.687	0.499	2.670	0.276
Florida	42	37	9.70 (0.525)	3.560	32.67	0.993	0.927	0.527	1.296	0.147
Pennsylvania	39	29	9.25 (0.770)	3.268	23.40	0.982	0.887	0.506	0.795	0.113
Maine	33	26	9.17 (0.820)	3.15	19.80	0.979	0.846	0.504	2.191	0.223
North Carolina	26	21	9.06 (0.841)	2.938	16.10	0.976	0.844	0.489	0.282	0.043

<sup>a</sup> Data were not clone corrected before calculations. Data are divided into 2011 to 2012 and 2013 to 2016 based on the simple sequence repeat primer panel used to acquire the data (9 loci versus 12 loci multiplex). eMLG, expected number of multilocus genotypes (MLG) at smallest size of at least 10; G, Stoddart and Taylor's index of MLG diversity; H, Shannon and Weiner index of multilocus genotype diversity; Hexp, Nei's 1978 gene diversity; Ia, index of association; MLG, number of multilocus genotypes; SE, standard error;  $\bar{r}_d$ , standardized index of association; N, number of individuals;  $\lambda$ , Simpson's index.

<sup>b</sup> Simpson's index was corrected for influence from sample size by multiplying  $\lambda$  by N/(N-1), in which N is the number of individuals.

Our data show that many of the tomato-specific lineages, including US-21, have the avirulent form of *PiAVR2*, whereas the potato lineages US-8, US-14, and US-24 have the virulent variant PiAVR2like. Because these three lineages most likely migrated from Mexico, they were probably exposed to the R2 resistance gene in Mexico, and most likely, the selection for the virulent variant occurred there. The tomato R genes, including the Ph-1, Ph-2, and *Ph-3* genes, are known to be responsible for host resistance in tomato, but their corresponding effectors have yet to be fully cloned and identified and were not evaluated in our work (Chunwongse et al. 2002; Hansen et al. 2014; Moreau et al. 1998; Nowicki et al. 2012; Vleeshouwers et al. 2011). Additional examination of a fuller set of effectors for host-specific patterns and additional tests to determine which ones may be under host selection in the field are needed, and they may give insight into host-specific adaptation in tomato.

Despite US-1's ability to infect both potato and tomato, the virulent variant *PiAVR2-like* was not detected in the isolates that we examined. Other studies have shown that US-1 has a narrower spectrum of virulence on R genes than other lineages (Goodwin et al. 1995b; Perez et al. 2001; Secor and Gudmestad 1999). However, in a study by Perez et al. (2001), US-1 in Peru was virulent on the R2 differentials that they examined. This suggests that within-lineage differences in the virulence spectrum of US-1 may have occurred.

**Population structure and source of U.S. lineages.** Examination of the genetic relatedness of the lineages by several different methods revealed consistent trends that suggest relationships between groups of U.S. lineages. US-1 shared little genetic similarity with subsequent lineages detected in the United States (Figs. 2, 3, and 4), suggesting that it may have migrated from a different source than other U.S. lineages. During the emergence of the mefenoxam-resistant lineages in the mid-1990s, US-1 rapidly declined (Goodwin et al. 1998). Also, because of the allelic differences at multiple SSR loci between US-1 and the other U.S. lineages and from our previously published whole-genome singlenucleotide polymorphism datasets (Martin et al. 2016), it is unlikely that US-1 was a parent for any subsequent lineages detected in US.

US-23 is currently the dominant lineage on both tomato and potato in the United States. Our data show considerable subclonal variation within the US-23 lineage. The MSNs from isolates collected over time clearly show the increase in the number of haplotypes within the US-23 lineage. In fact, the number of MLGs increased from 91 to 171 between 2011 and 2016. We also reported large amounts of subclonal variation within an aggressive Blue 13 lineage of *P. infestans* in Indian populations within a relatively short period of time after introduction (Dey et al. 2018). Thus, subclonal evolution can occur rapidly in the pathogen.

Potato-dominant lineages US-8, US-14, and US-24 exhibited strong levels of genetic similarity and clustered together in the MSN and neighbor joining trees. US-8 was among the first mefenoxamresistant lineages detected in the United States, and it has been proposed to be a putative parent to multiple subsequent lineages (Goodwin et al. 1998). US-14 was proposed to be a clonal derivative of US-8 (Goodwin et al. 1998). Both share a large number of similarities, including identical RFLP fingerprints and nearly identical SSR fingerprints, and they were positive for the PiAVR2like variant, supporting this idea. US-24 was first observed during the 2009 epidemic on potato in North Dakota (Fig. 1) (Hu et al. 2012). US-24 clustered with US-8 and US-14; it is also primarily a potato pathogen, but it is sensitive to mefenoxam. In addition, US-24 displays similar pathogenicity characteristics to US-8 (Danies et al. 2013), suggesting that US-24 may be a clonal derivative of US-8. US-8 has been continually detected in low frequencies (Fig. 1) (Saville et al. 2015). Our data showed US-8, US-14, and US-24 lineages on one branch of the MSN (Fig. 5). Multiple isolates of US-8, US-11, US-23, and US-24 were genotyped by sequencing, and a close relationship between US-8 and US-24 was also shown in that work (Hansen et al. 2016). US-8, US-14, and US-24 showed evidence of recombination in the *Piypt1* locus. Additional whole-genome sequencing of more lineages is needed to confirm or refute the clonal derivative hypothesis

U.S. lineages US-21 and US-22 shared some genetic similarity as indicated by the Structure and MSN analyses. These lineages formed a separate clade in the neighbor joining tree that was shared with one Mexican isolate (Fig. 3). Both lineages are primarily pathogenic to tomato and were first reported around the same time period in North Carolina and Florida (Hu et al. 2012; Schultz et al. 2010). Within the neighbor joining tree, the clade containing US-20 was nested within a larger clade consisting of other tomatopathogenic lineages, including US-12, US-13, and US-16. Within the DAPC, clusters for US-20, US-21, and US-22 displayed a closer proximity to the Mexican cluster than the South American clusters, suggesting a potential migration via Mexican tomatoes. Some Florida tomato growers also grow tomatoes in Mexico, and these fruit are shipped into the United States. The US-21 lineage has not been observed in the United States since 2011 and was only found in Florida and North Carolina on tomato, suggesting that it was not as fit as the US-23 lineage. US-21 was difficult to isolate in culture and very slow growing (J. B. Ristaino, *personal observations*). The US-22 lineage was widely prevalent during the 2009 U.S. pandemic but subsequently declined. However, the prevalence of US-22 during the 2009 epidemic was owing to human-mediated transport of infected transplants rather than through increased virulence (Fry et al. 2013).

The remaining U.S. lineages shared varying levels of genetic similarity among each other and among Mexican isolates. Only US-11 has been previously empirically shown to be a recombinant, and it is regarded as the first new recombinant to appear in the United States (Gavino et al. 2000; Goodwin et al. 1998). US-11 also showed evidence for recombination in our study based on examinations of the ancestral recombination events within *Piypt1*. All of the remaining lineages were derived from either newer or more distant recombination events. It is unclear if this was the result of recombination events in the United States or if they migrated into the United States from another source after the recombination event, most likely within a sexually reproducing population in Mexico. The latter explanation is more likely. US-11 as well as several rare lineages, such as US-12, US-15, and US-16, were positive for the avirulent PiAVR2 gene. With the exception of US-11, which continues to persist locally in the western United States (Fig. 1), all of these lineages have not been detected in the United States for many years.

We compared the structure of U.S. lineages of P. infestans with lineages from sources from South America and Mexico. Tomatoes are routinely shipped from Mexico and South America into the United States. Mexico has been one of the primary exporters of both field and greenhouse tomatoes to the United States for many years, with a dramatic increase in the export of greenhouse tomatoes in 2005 (Thornsbury and Jerardo 2012). Over 1.6 billion pounds of fresh greenhouse tomatoes were imported into the United States from Mexico in 2014 (Wells et al. 2015), providing a likely migration route for new lineages into the United States. The pathogen can survive in infected fruit and spread from culled fruit to neighboring fields near packing houses. We found shared genetic similarity and clustering between most U.S. lineages (with the exception of US-23 and US-1) and Mexican lineages. In the 1980s and 1990s, Mexico shipped infected tomatoes from the northwestern region of the country into the United States and continues to be a major supplier of U.S. tomatoes today (Fry 2008; Goodwin et al. 1998). One of these shipments was thought to have brought US-6 to the United States, whereas later shipments brought US-7 and US-8 (Goodwin 1997).

Multiple new lineages were observed during the mid-1990s, some of which were not observed again beyond their initial discovery. Four of these lineages, US-12, US-13, US-14, and US-15,

were all observed in the northeastern part of the United States. Danies et al. (2014) documented the presence of an ephemeral sexual population in the northeastern United States in 2010 and 2011. Given the large number of recorded outbreaks in the northeastern United States each year and the presence of unique variants (Fig. 1), it is a region that would be a likely area for the establishment of a future stable sexual population, and it should continue to be monitored closely.

Examination of Structure, neighbor joining, Hudson's statistics, and DAPC results showed that US-23 shared little similarity to other U.S. lineages. US-23 shared more genetic similarity with isolates from South America than with those from populations in central Mexico or other U.S. lineages. US-23 shared the greatest similarity with isolates identified as the BR-1 clonal lineage, which was first reported in the 1990s from outbreaks in Bolivia and Brazil. Others have reported that US-23 does occur in Mexico and shared genetic similarity with subpopulations from the Mexican state of Tlaxcalab. However, we did not examine subpopulations from Tlaxcalab in our work, because the datasets were not published when our work was done (Wang et al. 2017). The Mexican isolates used in our work were from the Toluca region. Our analysis indicates that US-23 shares genetic similarity with BR-1 populations from South America rather than the central Mexican populations that we examined. The US-23 and BR-1 lineages were not identical but genetically similar sister lineages in our neighbor joining tree and DAPC (Fig. 3). Our data suggest that US-23 did not arrive from the same Mexican source that we studied as the other U.S. lineages. If South America was a source, importation of infected fruit, tubers, or plant material from that region is a likely source.

Examination of potential migration scenarios indicated that the scenario with the highest probability was that US-23 diverged from a South American source, further indicating that the migration source of US-23 was not the same as many of the other U.S. lineages. The US-23 lineage has been reported in Europe since 2012, and it has been named the 23\_A1 (Kröner et al. 2017; Stroud et al. 2016). Additional analysis with more European populations might document the genetic relationship between BR-1, US-23, the European 23\_A1, and the Tlaxcalab Mexican populations. That work is planned.

Although US-23 shares the most genetic similarity with isolates of BR-1 from Bolivia and Brazil, it is unlikely that US-23 was directly imported to the United States on tomato fruit from either country. Both countries have been excluded from exportation of fresh tomatoes to the United States since February 2009 because of concerns over potential introduction of the tomato leaf miner (Tuta absoluta Meyrick) (Bech 2009). However, BR-1 is predominantly a potato lineage, and importation of potato tubers from Brazil is not restricted. The BR-1 lineage has also been found in other countries in South America, including Argentina and Uruguay (Adler et al. 2002), and other sources of inoculum are likely, including through plant material exchange via trade or breeding programs. The BR-1 lineage may have been introduced into Bolivia and Brazil from Europe. European potatoes have been shipped into both Bolivia and Brazil in the past, and the possibility that BR-1 migrated to South America from Europe should be explored further. Other European lineages have been found in Argentina and Chile, including the new lineage EU\_2\_A1 that is currently widespread in several South American countries (Ristaino et al. 2019).

US-23 has become the dominant lineage in the United States, replacing previously dominant lineages, including US-8 and US-22 (Saville et al. 2015). US-23 is also mefenoxam sensitive (Danies et al. 2013; Saville et al. 2015). Studies to determine why US-23 has displaced other lineages suggest that it may be at least in part because of greater infection efficiency when compared with other contemporary lineages (Fall et al. 2015). A study of acquired mefenoxam resistance in *P. infestans* has also suggested that there may be a fitness cost associated with fungicide resistance (Childers et al. 2015). US-23 has also been shown to be more resilient to

overwintering conditions in host tissue, and it survives longer with less loss in growth than US-22 or US-24, allowing it to not only overwinter but outgrow other lineages (Frost et al. 2016). In addition, US-23 has greater sporulation rates than US-22 or US-24 and greater epidemic potential based on computer simulations of outbreaks, suggesting that US-23 has a greater overall fitness and allowing it to currently dominate U.S. populations (Seidl Johnson et al. 2015).

**Population structure of U.S. lineages, 2011 to 2016.** *P. infestans* can move within the United States from several known migration paths, including (i) Florida northward from winter tomatoes and potatoes; (ii) from seed potato-producing areas in the western United States, such as North Dakota; and (iii) from seed potatoes from Maine and Canada into the eastern United States. These migration routes may be the result of long-distance shipment of infected seed potatoes or via infected tomato fruit. US-11, which has been detected in low frequencies in the western United States, is also prevalent in Taiwan, suggesting the movement of the pathogen from Asia. The movement of the pathogen from Florida northward can occur in infected tomato fruit and potatoes that are repacked and subsequently culled at packing houses. A more detailed examination of the movement of material from Florida and the testing of cull piles may provide more insight as to the nature of this movement.

The composition and number of lineages observed in the United States have undergone extensive changes over the past 7 years. From 2011 to 2016, the total number of clonal lineages observed decreased, and the dominance of the US-23 lineage occurred. The number of MLGs identified as US-23 increased rapidly over the same time period. Previous work by Hansen et al. (2016) using genotype by sequencing showed multiple clades and subclonal variation of US-23, suggesting genetic diversification within the lineage. Our data support this conclusion, because the number of MLGs expanded greatly over the time period examined.

Observation of the samples by geographic region revealed differences in lineage composition depending on whether the sample was collected on the East Coast or the West Coast of the United States (Fig. 1A). Samples collected from the West Coast were composed primarily of lineages more rarely observed, such as US-8 and US-11, and showed little subclonal variation. By comparison, samples from the East Coast were mostly US-23 and had a much larger number of MLGs and large amounts of subclonal variation. The difference in genotypic composition suggests differences in inoculum sources and suggests the likely survival of US-8 in West Coast potato seed sources.

Patterns in the genetic structure of U.S. lineages suggest that the changes in the composition of lineages over the past 30 years have been primarily the result of multiple introductions of P. infestans into the United States, most likely from Mexico. The USABlight database needs to be expanded to include disease reports and genotyping of lineages from Mexico and Canada. We are in the process of expanding the maps to include these countries. The United States also imports fresh market tomatoes from The Netherlands, and therefore, additional comparison with European genotypes of the pathogen are needed. A shared global database of P. infestans lineages that includes U.S., European, Mexican, and South American lineages genotyped by SSR lineage to identify and track spread is greatly needed. Currently, databases are fragmented, and they are not fully functional for examining global outbreak datasets (Ristaino et al. 2019). Our data suggest that enhanced phytosanitary inspections of imported fruit and planting material particularly from Mexico and use of molecular diagnostics at ports of entry and quarantine may be needed to prevent introduction of new virulent lineages of *P. infestans* into the United States.

**Data accessibility.** Sequences of representative haplotypes for each locus studied were deposited in the NCBI GenBank (*Piypt:* accession numbers MG010330 to MG010343; *PiAVR2:* accession number MG010344; and *PiAVR2-like:* accession numbers MG010345 to MG010346). SSR lineages are deposited in USABlight.org.

#### ACKNOWLEDGMENTS

We thank William Fry and Kevin Myers at Cornell University for sharing SSR datasets from USABlight. Data are archived on USABlight.org. The authors also thank the many USABlight researchers (K. Everts, B. Fry, A. Gevens, N. Grünwald, B. Gugino, S. Johnson, D. Johnson, H. Judelson, M. McGrath, P. Roberts, C. Smart, and G. Secor) and international collaborators (including Eduardo Mitzibuti, Greg Forbes, Enrique Northcote, Wilmer Perez, Soledad Gamboa CIP Peru, and ProINPA, Bolivia) that have provided cultures over many years or plant samples for genotyping. We thank the National Institute of Food and Agriculture interns Meghan Wyatt, Sarah Turner, Kenneth Lee, and Rachael Guenter for technical assistance. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### LITERATURE CITED

- Adler, N., Chacón, G., Forbes, G., and Flier, W. 2002. *Phytophthora infestans* sensu lato in South America population substructuring through host-specificity. Pages 13-17 in: Late Blight: Managing the Global Threat: Proceedings of the Global Initiative on Late Blight Conference. C. Lizárraga, ed. International Potato Center, (CIP), Lima, Peru.
- Aylor, D. L., Price, E. W., and Carbone, I. 2006. SNAP: Combine and Map modules for multilocus population genetic analysis. Bioinformatics 22: 1399-1401.
- Bech, R. A. 2009. Federal order for tomatoes from countries where *Tuta absoluta* is known to occur. Animal and Plant Health Inspection Service: Plant Protection and Quarantine Report. USDA APHIS, Washington DC.
- Bourke, P. M. A. 1964. Emergence of potato blight, 1843-46. Nature 203: 805-808.
- Bowden, L. C., Price, E. W., and Carbone, I. 2008. SNAP Clade and Matrix, Version 2.
- Brown, A. H. D., Feldman, M. W., and Nevo, E. 1980. Multilocus structure of natural populations of *Hordeum spontaneum*. Genetics 96:523-526.
- Brurberg, M. B., Elameen, A., Le, V. H., Nærstad, R., Hermansen, A., Lehtinen, A., Hannukkala, A., Nielsen, B., Hansen, J., Andersson, B., and Yuen, J. 2011. Genetic analysis of *Phytophthora infestans* populations in the Nordic European countries reveals high genetic variability. Fungal Biol. 115:335-342.
- Cárdenas, M., Grajales, A., Sierra, R., Rojas, A., González-Almario, A., Vargas, A., Marín, M., Fermín, G., Lagos, L. E., Grünwald, N. J., Bernal, A., Salazar, C., and Restrepo, S. 2011. Genetic diversity of *Phytophthora infestans* in the Northern Andean region. BMC Genet. 12:23.
- Childers, R., Danies, G., Myers, K., Fei, Z., Small, I. M., and Fry, W. E. 2015. Acquired resistance to mefenoxam in sensitive isolates of *Phytophthora infestans*. Phytopathology 105:342-349.
- Chunwongse, J., Chunwongse, C., Black, L., and Hanson, P. 2002. Molecular mapping of the *Ph-3* gene for late blight resistance in tomato. J. Hortic. Sci. Biotechnol. 77:281-286.
- Cornuet, J. M., Ravigne, V., and Estoup, A. 2010. Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). BMC Bioinformatics 11:401.
- Danies, G., Myers, K., Mideros, M., Restrepo, S., Martin, F. N., Cooke, D. E. L., Smart, C. D., Ristaino, J. B., Seaman, A. J., Gugino, B. K., Grünwald, N. J., and Fry, W. E. 2014. An ephemeral sexual population of *Phytophthora infestans* in the northeastern United States and Canada. PLoS One 9:e116354.
- Danies, G., Small, I. M., Myers, K., Childers, R., and Fry, W. E. 2013. Phenotypic characterization of recent clonal lineages of *Phytophthora infestans* in the United States. Plant Dis. 97:873-881.
- Dey, T., Saville, A., Myers, K., Tewan, S., Cooke, D. E. L., Tripathy, S., Fry, W. E., Ristaino, J. B., and Roy, S. G. 2018. Large sub-clonal variation in *Phytophthora infestans* from recent severe late blight epidemics in India. Sci. Rep. 8:4429-4441.
- Drenth, A., Goodwin, S., Fry, W., and Davidse, L. 1993. Genotypic diversity of *Phytophthora infestans* in the Netherlands revealed by DNA polymorphisms. Phytopathology 83:1087-1092.
- Earl, D. A., and vonHoldt, B. M. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4: 359-361.
- Excoffier, L., and Lischer, H. E. L. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. 10:564-567.
- Fall, M. L., Tremblay, D. M., Gobeil-Richard, M., Couillard, J., Rocheleau, H., Van der Heyden, H., Levesque, C. A., Beaulieu, C., and Carisse, O. 2015. Infection efficiency of four *Phytophthora infestans* clonal lineages and DNA-based quantification of sporangia. PLoS One 10:e0136312.

- Flot, J. F. 2010. SeqPHASE: A web tool for interconverting phase input/output files and fasta sequence alignments. Mol. Ecol. Resour. 10:162-166.
- Forbes, G. A., Goodwin, S. B., Drenth, A., Oyarzun, P., Ordoñez, M. E., and Fry, W. E. 1998. A global marker database for *Phytophthora infestans*. Plant Dis. 82:811-818.
- Frost, K. E., Seidl Johnson, A. C., and Gevens, A. J. 2016. Survival of isolates of the US-22, US-23, and US-24 clonal lineages of *Phytophthora infestans* by asexual means in tomato seed at cold temperatures. Plant Dis. 100: 180-187.
- Fry, W. 2008. *Phytophthora infestans*: The plant (and R gene) destroyer. Mol. Plant Pathol. 9:385-402.
- Fry, W. E., Birch, P. R. J., Judelson, H. S., Grünwald, N. J., Danies, G., Everts, K. L., Gevens, A. J., Gugino, B. K., Johnson, D. A., Johnson, S. B., McGrath, M. T., Myers, K. L., Ristaino, J. B., Roberts, P. D., Secor, G., and Smart, C. D. 2015. Five reasons to consider *Phytophthora infestans* a reemerging pathogen. Phytopathology 105:966-981.
- Fry, W. E., and Goodwin, S. B. 1997. Re-emergence of potato and tomato late blight in the United States. Plant Dis. 81:1349-1357.
- Fry, W. E., Myers, K., Roberts, P. D., McGrath, M. T., Everts, K. L., Secor, G., Seaman, A., Gevens, A. J., Seebold, K., Jr., Zitter, T. A., Gugino, B. K., Snover-Clift, K., McLeod, A., Johnson, S. B., Wyenandt, A., Danies, G., Judelson, H. S., Grünwald, N., Small, I. M., Ristaino, J., and Smart, C. D. 2013. The 2009 late blight pandemic in the Eastern United States - causes and results. Plant Dis. 97:296-306.
- Fu, Y. X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147:915-925.
- Garry, G., Forbes, G. A., Salas, A., Santa Cruz, M., Perez, W. G., and Nelson, R. J. 2005. Genetic diversity and host differentiation among isolates of *Phytophthora infestans* from cultivated potato and wild solanaceous hosts in Peru. Plant Pathol. 54:740-748.
- Gavino, P. D., Smart, C. D., Sandrock, R. W., Miller, J. S., Hamm, P. B., Lee, T. Y., Davis, R. M., and Fry, W. E. 2000. Implications of sexual reproduction for *Phytophthora infestans* in the United States: Generation of an aggressive lineage. Plant Dis. 84:731-735.
- Gilroy, E. M., Breen, S., Whisson, S. C., Squires, J., Hein, I., Kaczmarek, M., Turnbull, D., Boevink, P. C., Lokossou, A., Cano, L. M., Morales, J., Avrova, A. O., Pritchard, L., Randall, E., Lees, A., Govers, F., van West, P., Kamoun, S., Vleeshouwers, V. G., Cooke, D. E., and Birch, P. R. 2011. Presence/absence, differential expression and sequence polymorphisms between *PiAVR2* and *PiAVR2-like* in *Phytophthora infestans* determine virulence on R2 plants. New Phytol. 191:763-776.
- Gómez-Alpizar, L., Carbone, I., and Ristaino, J. B. 2007. An Andean origin of *Phytophthora infestans* inferred from mitochondrial and nuclear gene genealogies. Proc. Natl. Acad. Sci. USA 104:3306-3311.
- Goodwin, S. B. 1997. The population genetics of *Phytophthora*. Phytopathology 87:462-473.
- Goodwin, S. B., Smart, C. D., Sandrock, R. W., Deahl, K. L., Punja, Z. K., and Fry, W. E. 1998. Genetic change within populations of *Phytophthora infestans* in the United States and Canada during 1994 to 1996: Role of migration and recombination. Phytopathology 88:939-949.
- Goodwin, S. B., Sujkowski, L. S., Dyer, A. T., Fry, B. A., and Fry, W. E. 1995a. Direct detection of gene flow and probable sexual reproduction of *Phytophthora infestans* in northern North America. Phytopathology 85: 473-479.
- Goodwin, S. B., Sujkowski, L. S., and Fry, W. E. 1995b. Rapid evolution of pathogenicity within clonal lineages of the potato late blight disease fungus. Phytopathology 85:669-676.
- Goodwin, S. B., Sujkowski, L. S., and Fry, W. E. 1996. Widespread distribution and probable origin of resistance to metalxyl in clonal genotypes of *Phytophthora infestans* in the United States and western Canada. Phytopathology 86:793-800.
- Grünwald, N. J., Goodwin, S. B., Milgroom, M. G., and Fry, W. E. 2003. Analysis of genotypic diversity data for populations of microorganisms. Phytopathology 93:738-746.
- Hall, T. A. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/08/NT. Nucleic Acids Symp. Ser. 41:95-98.
- Hansen, Z. R., Everts, K. L., Fry, W. E., Gevens, A. J., Grünwald, N. J., Gugino, B. K., Johnson, D. A., Johnson, S. B., Judelson, H. S., Knaus, B. J., McGrath, M. T., Myers, K. L., Ristaino, J. B., Roberts, P. D., Secor, G. A., and Smart, C. D. 2016. Genetic variation within clonal lineages of *Phytophthora infestans* revealed through genotyping-by-sequencing, and implications for late blight epidemiology. PLoS One 11:e0165690.
- Hansen, Z. R., Small, I. M., Mutschler, M., Fry, W. E., and Smart, C. D. 2014. Differential susceptibility of 39 tomato varieties to *Phytophthora infestans* clonal lineage US-23. Plant Dis. 98:1666-1670.
- Hu, C. H., Perez, F. G., Donahoo, R., McLeod, A., Myers, K., Ivors, K., Secor, G., Roberts, P. D., Deahl, K. L., Fry, W. E., and Ristaino, J. B. 2012. Recent genotypes of *Phytophthora infestans* in the eastern United States reveal

clonal populations and reappearance of mefenoxam sensitivity. Plant Dis. 96:1323-1330.

- Hudson, R. R., Boos, D. D., and Kaplan, N. L. 1992a. A statistical test for detecting geographic subdivision. Mol. Biol. Evol. 9:138-151.
- Hudson, R. R., Slatkin, M., and Maddison, W. P. 1992b. Estimation of levels of gene flow from DNA sequence data. Genetics 132:583-589.
- Hudson, R. R. 2000. A new statistic for detecting genetic differentiation. Genetics 155:2011-2014.
- Jakobsson, M., and Rosenberg, N. A. 2007. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801-1806.
- Jombart, T. 2008. adegenet: A R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403-1405.
- 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.
- Knaus, B. J., Tabima, J. F., Davis, C. E., Judelson, H. S., and Grünwald, N. J. 2016. Genomic analyses of dominant U.S. clonal lineages of *Phytophthora infestans* reveals a shared common ancestry for clonal lineages US11 and US18 and a lack of recently shared ancestry among all other U.S. lineages. Phytopathology 106:1393-1403.
- Kröner, A., Mabon, R., Corbière, R., Montarry, J., and Andrivon, D. 2017. The coexistence of generalist and specialist clonal lineages in natural populations of the Irish famine pathogen *Phytophthora infestans* explains local adaptation to potato and tomato. Mol. Ecol. 26:1891-1901.
- Lees, A., Wattier, R., Shaw, D. S., Sullivan, L., Williams, N., and Cooke, D. E. 2006. Novel microsatellite markers for the analysis of *Phytophthora infestans* populations. Plant Pathol. 55:311-319.
- Li, Y., Cooke, D. E. L., Jacobsen, E., and van der Lee, T. 2013. Efficient multiplex simple sequence repeat genotyping of the oomycete plant pathogen *Phytophthora infestans*. J. Microbiol. Methods 92:316-322.
- Ludwig, J. A. R. 1988. Statistical Ecology: A Primer on Methods and Computing. Wiley, New York, NY.
- Lyngsø, R. B., Song, Y. S., and Hein, J. 2005. Minimum recombination histories by branch and bound. Pages 239-250 in: Algorithms in Bioinformatics: Proceedings of the 5th International Workshop, WABI 2005. R. Casadio and G. Myers, eds. Springer, Berlin, Germany.
- Martin, M. D., Cappellini, E., Samaniego, J. A., Zepeda, M. L., Campos, P. F., Seguin-Orlando, A., Wales, N., Orlando, L., Ho, S. Y., Dietrich, F. S., Mieczkowski, P. A., Heitman, J., Willerslev, E., Krogh, A., Ristaino, J. B., and Gilbert, M. T. 2013. Reconstructing genome evolution in historic samples of the Irish potato famine pathogen. Nat. Commun. 4:2172.
- Martin, M. D., Ho, S. Y. W., Wales, N., Ristaino, J. B., and Gilbert, M. T. P. 2014. Persistence of the mitochondrial lineage responsible for the Irish potato famine in extant New World *Phytophthora infestans*. Mol. Biol. Evol. 31:1414-1420.
- Martin, M. D., Vieira, F. G., Ho, S. Y., Wales, N., Schubert, M., Seguin-Orlando, A., Ristaino, J. B., and Gilbert, M. T. 2016. Genomic characterization of a South American *Phytophthora* hybrid mandates reassessment of the geographic origins of *Phytophthora infestans*. Mol. Biol. Evol. 33:478-491.
- Matson, M. E. H., Small, I. M., Fry, W. E., and Judelson, H. S. 2015. Metalaxyl resistance in *Phytophthora infestans*: Assessing role of RPA190 gene and diversity within clonal lineages. Phytopathology 105:1594-1600.
- May, K. J., and Ristaino, J. B. 2004. Identity of the mtDNA haplotype(s) of *Phytophthora infestans* in historical specimens from the Irish potato famine. Mycol. Res. 108:471-479.
- Monacell, J. T., and Carbone, I. 2014. Mobyle SNAP Workbench: A webbased analysis portal for population genetics and evolutionary genomics. Bioinformatics 30:1488-1490.
- Montarry, J., Hamelin, F. M., Glais, I., Corbière, R., and Andrivon, D. 2010. Fitness costs associated with unnecessary virulence factors and life history traits: Evolutionary insights from the potato late blight pathogen *Phytophthora infestans*. BMC Evol. Biol. 10:283.
- Moreau, P., Thoquet, P., Olivier, J., Laterrot, H., and Grimsley, N. 1998. Genetic mapping of *Ph-2*, a single locus controlling partial resistance to *Phytophthora infestans* in tomato. Mol. Plant-Microbe Interact. 11: 259-269.
- Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583-590.
- Nowicki, M., Foolad, M. R., Nowakowska, M., and Kozik, E. U. 2012. Potato and tomato late blight caused by *Phytophthora infestans*: An overview of pathology and resistance breeding. Plant Dis. 96:4-17.
- Oyarzun, P. J., Pozo, A., Ordoñez, M. E., Doucett, K., and Forbes, G. A. 1998. Host specificity of *Phytophthora infestans* on tomato and potato in Ecuador. Phytopathology 88:265-271.
- Paradis, E. 2010. pegas: An R package for population genetics with an integrated-modular approach. Bioinformatics 26:419-420.

- Perez, W. G., Gamboa, J. S., Falcon, Y. V., Coca, M., Raymundo, R. M., and Nelson, R. J. 2001. Genetic Structure of Peruvian Populations of *Phytophthora infestans*. Phytopathology 91:956-965.
- Pielou, E. C. 1975. Ecological Diversity. Wiley, New York.
- Pritchard, J. K., Stephens, M., and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.
- Ristaino, J., Cooke, D. E., Acuña, I., and Muñoz, M. 2019. The threat of late blight to global food security. In: Emerging Plant Disease and Global Food Security. A. Records and J. Ristaino, eds. American Phytopathological Society Press, St. Paul, MN.
- Rosenberg, N. A. 2004. DISTRUCT: A program for the graphical display of population structure. Mol. Ecol. Notes 4:137-138.
- Saville, A., Graham, K., Grünwald, N. J., Kevin, M., Fry, W. E., and Ristaino, J. B. 2015. Fungicide sensitivity of U.S. genotypes of *Phytophthora infestans* to six oomycete-targeted compounds. Plant Dis. 99:659-666.
- Saville, A. C., Martin, M. D., and Ristaino, J. B. 2016. Historic late blight outbreaks caused by a widespread dominant lineage of *Phytophthora infestans* (Mont.) de Bary. PLoS One 11:e0168381.
- Schultz, D., Donahoo, R., Perez, F. G., Tejeda, S., Roberts, P. D., and Deahl, K. L. 2010. A survey of tomato and potato fields in Florida reveals unique genotypes of *Phytophthora infestans* between 2005 and 2007. HortScience 45:1064-1068.
- Secor, G. A., and Gudmestad, N. C. 1999. Managing fungal diseases of potato. Can. J. Plant Pathol. 21:213-221.
- Seidl Johnson, A. C., Frost, K. E., Rouse, D. I., and Gevens, A. J. 2015. Effect of temperature on growth and sporulation of US-22, US-23, and US-24 clonal lineages of *Phytophthora infestans* and implications for late blight epidemiology. Phytopathology 105:449-459.
- Shannon, C. E. 2001. A mathematical theory of communication. SIGMOBILE Mob. Comput. Commun. Rev. 5:3-55.
- Simpson, E. H. 1949. Measurement of diversity. Nature 163:688.
- Smith, J. M., Smith, N. H., O'Rourke, M., and Spratt, B. G. 1993. How clonal are bacteria? Proc. Natl. Acad. Sci. USA 90:4384-4388.
- Stephens, M., and Donnelly, P. 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am. J. Hum. Genet. 73:1162-1169.
- Stoddart, J. A., and Taylor, J. F. 1988. Genotypic diversity: Estimation and prediction in samples. Genetics 118:705.
- Stroud, J. A., Shaw, D. S., Hale, M. D., and Steele, K. A. 2016. SSR assessment of *Phytophthora infestans* populations on tomato and potato in British gardens demonstrates high diversity but no evidence for host specialization. Plant Pathol. 65:334-341.
- Tajima, F. 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics 105:437-460.
- Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585-595.
- Teschemacher, J. E. 1845. Observations on the potato disease. Page 125 in: Gardeners' Chronicle and Agricultural Gazette. London, England.
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. 1994. Clustal-W-Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673-4680.
- Thornsbury, S., and Jerardo, A. 2012. Vegetables and Pulses Outlook. U.S. Department of Agriculture: Economic Research Service, Washington DC.
- Vleeshouwers, V. G. A. A., Raffaele, S., Vossen, J. H., Champouret, N., Oliva, R., Segretin, M. E., Rietman, H., Cano, L. M., Lokossou, A., Kessel, G., Pel, M. A., and Kamoun, S. 2011. Understanding and exploiting late blight resistance in the age of effectors. Annu. Rev. Phytopathol. 49:507-531.
- Wang, H., Qi, M. Q., and Cutler, A. J. 1993. A simple method of preparing plant samples for PCR. Nucleic Acids Res. 21:4153-4154.
- Wang, J., Fernández-Pavía, S. P., Larsen, M. M., Garay-Serrano, E., Gregorio-Cipriano, R., Rodríguez-Alvarado, G., Grünwald, N. J., and Goss, E. M. 2017. High levels of diversity and population structure in the potato late blight pathogen at the Mexico centre of origin. Mol. Ecol. 26:1091-1107.
- Wangsomboondee, T., Groves, C. T., Shoemaker, P. B., Cubeta, M. A., and Ristaino, J. B. 2002. *Phytophthora infestans* populations from tomato and potato in North Carolina differ in genetic diversity and structure. Phytopathology 92:1189-1195.
- Watterson, G. A. 1975. On the number of segregating sites in genetical models without recombination. Theor. Popul. Biol. 7:256-276.
- Wells, H. F., Bond, J., and Thornsbury, S. 2015. Vegetables and Pulses Outlook. U.S. Department of Agriculture: Economic Research Service, Washington DC.
- Yoshida, K., Schuenemann, V. J., Cano, L. M., Pais, M., Mishra, B., Sharma, R., Lanz, C., Martin, F. N., Kamoun, S., Krause, J., Thines, M., Weigel, D., and Burbano, H. A. 2013. The rise and fall of the *Phytophthora infestans* lineage that triggered the Irish potato famine. eLife 2:e00731.
- Yuen, J. E., and Andersson, B. 2013. What is the evidence for sexual reproduction of *Phytophthora infestans* in Europe? Plant Pathol. 62:485-491.