Title

Emergence of a Plant Pathogen in Europe Associated with Multiple Intercontinental Introductions.

Journal

Applied and Environmental Microbiology, 86(3)

ISSN

0099-2240

Authors

Landa, Blanca B
Castillo, Andreina I
Giampetruzzi, Annalisa
et al.

Publication Date

2020-01-21

DOI

10.1128/aem.01521-19

Peer reviewed

AMERICAN SOCIETY FOR MICROBIOLOGY | Applied and Environmental Microbiology®

EVOLUTIONARY AND GENOMIC MICROBIOLOGY

Emergence of a Plant Pathogen in Europe Associated with Multiple Intercontinental Introductions

Blanca B. Landa,[a] Andreina I. Castillo,[b] Annalisa Giampetruzzi,[c] Alexandra Kahn,[b] Miguel Román-Écija,[a] María Pilar Velasco-Amo,[a] Juan A. Navas-Cortés,[a] Ester Marco-Noales,[d] Silvia Barbé,[d] Eduardo Moralejo,[e] Helvecio D. Coletta-Filho,[f] Pasquale Saldarelli,[g] Maria Saponari,[g] Rodrigo P. P. Almeida[b]

[a]Institute for Sustainable Agriculture, Consejo Superior de Investigaciones Científicas (IAS-CSIC), Córdoba, Spain
[b]Department of Environmental Science, Policy and Management, University of California, Berkeley, Berkeley, California, USA
[c]Dipartimento di Scienze del Suolo della Pianta e degli Alimenti, Universit à degli Studi di Bari Aldo Moro, Bari, Italy
[d]Centro de Protección Vegetal y Biotecnología, Instituto Valenciano de Investigaciones Agrarias (IVIA), Moncada, Spain
[e]Tragsa, Empresa de Transformación Agraria, Delegación de Baleares, Palma de Mallorca, Spain
[f]Centro de Citricultura Sylvio Moreira (IAC), Cordeirópolis, São Paulo, Brazil
[g]Istituto per la Protezione Sostenibile delle Piante, CNR, Bari, Italy

Blanca B. Landa, Andreina I. Castillo, and Annalisa Giampetruzzi contributed equally to this work.

ABSTRACT Pathogen introductions have led to numerous disease outbreaks in naive regions of the globe. The plant pathogen *Xylella fastidiosa* has been associated with various recent epidemics in Europe affecting agricultural crops, such as almond, grapevine, and olive, but also endemic species occurring in natural forest landscapes and ornamental plants. We compared whole-genome sequences of *X. fastidiosa* subspecies *multiplex* from America and strains associated with recent outbreaks in southern Europe to infer their likely origins and paths of introduction within and between the two continents. Phylogenetic analyses indicated multiple introductions of *X. fastidiosa* subspecies *multiplex* into Italy, Spain, and France, most of which emerged from a clade with limited genetic diversity with a likely origin in California, USA. The limited genetic diversity observed in *X. fastidiosa* subspecies *multiplex* strains originating from California is likely due to the clade itself being an introduction from *X. fastidiosa* subspecies *multiplex* populations in the southeastern United States, where this subspecies is most likely endemic. Despite the genetic diversity found in some areas in Europe, there was no clear evidence of recombination occurring among introduced *X. fastidiosa* strains in Europe. Sequence type taxonomy, based on multilocus sequence typing (MLST), was shown, at least in one case, to not lead to monophyletic clades of this pathogen; whole-genome sequence data were more informative in resolving the history of introductions than MLST data. Although additional data are necessary to carefully tease out the paths of these recent dispersal events, our results indicate that whole-genome sequence data should be considered when developing management strategies for *X. fastidiosa* outbreaks.

IMPORTANCE *Xylella fastidiosa* is an economically important plant-pathogenic bacterium that has emerged as a pathogen of global importance associated with a devastating epidemic in olive trees in Italy associated with *X. fastidiosa* subspecies *pauca* and other outbreaks in Europe, such as *X. fastidiosa* subspecies *fastidiosa* and *X. fastidiosa* subspecies *multiplex* in Spain and *X. fastidiosa* subspecies *multiplex* in France. We present evidence of multiple introductions of *X. fastidiosa* subspecies *multiplex*, likely from the United States, into Spain, Italy, and France. These introductions illustrate the risks associated with the commercial trade of plant material at global scales and the need to develop effective policy to limit the likelihood of pathogen pollution into naive regions. Our study demonstrates the need to utilize

Citation Landa BB, Castillo AI, Giampetruzzi A, Kahn A, Román-Écija M, Velasco-Amo MP, Navas-Cortés JA, Marco-Noales E, Barbé S, Moralejo E, Coletta-Filho HD, Saldarelli P, Saponari M, Almeida RPP. 2020. Emergence of a plant pathogen in Europe associated with multiple intercontinental introductions. Appl Environ Microbiol 86:e01521-19. https://doi.org/10.1128/AEM.01521-19.

Editor Emma R. Master, University of Toronto

Address correspondence to Blanca B. Landa, blanca.landa@csic.es.

Received 12 July 2019
Accepted 5 November 2019
Accepted manuscript posted online 8 November 2019
Published 21 January 2020

whole-genome sequence data to study *X. fastidiosa* introductions at outbreak stages, since a limited number of genetic markers does not provide sufficient phylogenetic resolution to determine dispersal paths or relationships among strains that are of biological and quarantine relevance.

KEYWORDS *Xylella fastidiosa*, emerging disease, genomic diversity, outbreaks, recombination, quarantine

The risks associated with emerging infectious diseases to humans, animals, and plants, have become increasingly recognized due to high-profile epidemics in recent years (1). The latest Ebola virus epidemics in Africa have highlighted the regional and global threats of zoonotic spillover events originating from pathogen endemic disease cycles (2). Likewise, the commercial trade of amphibian species, which facilitates and speeds the spread of pathogens at the global scale (3), contributed to massive amphibian decline due to infection with the chytrid fungus *Batrachochytrium dendrobatidis* (4). The chytrid-driven pandemic represents a case of pathogen pollution where anthropogenic activities, such as animal trade, led to movement of the fungus beyond its natural range (3, 5). Pathogen pollution has also been identified as a major contributor to novel plant diseases, with introductions being associated with over half of the emerging diseases reported (6). One recent emerging plant disease driven by such a mechanism is the ongoing epidemic of olive quick decline syndrome (OQDS) in southern Italy, caused by the introduction of a strain of the bacterium *Xylella fastidiosa* subsp. *pauca* from Central America into the region, likely via the commercial trade of infected plant material (7). Recently, *X. fastidiosa* has been intercepted in nursery plant material at European ports of entry (e.g., 8), and various naturally existing strains have also been reported in southern Europe, from Portugal to Italy (9–11). Understanding the mechanisms and paths of dispersal has become of significant importance, as introductions appear to continue and over 500 plant species are already listed as susceptible to infection by *X. fastidiosa* (12, 13).

Xylella fastidiosa is a xylem-limited plant pathogen spread by xylem sap-feeding insects. It is taxonomically divided into three major subspecies (*X. fastidiosa* subspp. *fastidiosa*, *multiplex*, and *pauca*) (14), although additional subspecies have been suggested (i.e., *X. fastidiosa* subspp. *sandyi* and *morus* [15]). The main subspecies have allopatric origins in the Americas but have dispersed and established various times, with North, Central, and South America serving as both sources and sinks of introductions (16–18). In addition to the subspecies categorization, *X. fastidiosa* can be subdivided into sequence types (STs) using a multilocus sequence typing (MLST) approach based on seven housekeeping genes (19). MLST is a portable and robust platform that is widely used to assign strains to genetic groups (20) while providing information on host plant species that may be susceptible to particular strains (21).

Xylella fastidiosa emerged as a pathogen of global importance in 2013, when it was found associated with olive trees in Italy, and since then it has been reported in several countries outside the Americas (7). Four subspecies of the pathogen (*X. fastidiosa* subspp. *fastidiosa*, *multiplex*, *sandyi*, and *pauca*) have been reported in Europe; moreover, up to nine STs, mostly of *X. fastidiosa* subsp. *multiplex*, have already been detected (12, 13). The taxonomic placement of *X. fastidiosa* strains in Europe is of applied relevance, as European Commission-mandated management strategies are based on the subspecies present in each outbreak. Nonetheless, the generalization of biological and ecological features of strains belonging to a subspecies is limited since fundamental aspects of *X. fastidiosa* biology, including plant host range, vary even at the more resolved ST level (e.g., 21). In addition, ST-level genetic resolution among strains may not permit the identification of closely related but different strains. Finally, *X. fastidiosa* is naturally competent (22), and while MLST is particularly powerful for the taxonomy of recombinogenic taxa, it is also possible that strains belonging to the same ST may not be phylogenetically related or share similar traits, such as host plant range. For

these reasons, whole-genome sequence data should probably be used when considering *X. fastidiosa* quarantine, eradication, and containment strategies.

Here, we focus on *X. fastidiosa* subsp. *multiplex*, the causal agent of several diseases in almond, oak, plum, and shade trees, among other plant species. Whole-genome sequence data were used to analyze the origin of *X. fastidiosa* subspecies *multiplex* introductions into Europe and Brazil. While some of the *X. fastidiosa* subsp. *multiplex* strains found in Europe have been previously described (i.e., CFBP8416, CFBP8417, and CFBP8418 from Corsica, France), the present study includes new isolates from Spain (Alicante and the Mallorca and Menorca islands) and Tuscany, Italy. Thus, the results provide a clearer picture of the history of *X. fastidiosa* subsp. *multiplex* introductions, likely from the United States, into Spain, Italy, and France, as well as into Brazil. We also investigated recombination patterns among European strains. The data also show that some STs may not be monophyletic. Finally, *X. fastidiosa* subspecies *multiplex* is divided into two groups based on MLST, one with limited and another with significant evidence of recombination (23). We found that this subdivision is supported by core genome sequence data and that *X. fastidiosa* subspecies *multiplex* strains obtained from European samples belong to the "non-IHR" group (isolates not showing intersubspecific homologous recombination), previously described by Nunney et al. (23). Moreover, our data indicate that the clade with limited recombination is relatively young, potentially explaining the difference in recombination rates. In summary, whole-genome sequences provide higher phylogenetic resolution than other genotyping methods and allow for better inferences on pathogen-spread pathways and mechanisms.

RESULTS

Multiple introductions of *X. fastidiosa* subsp. *multiplex* into Europe. A maximum likelihood (ML) phylogenetic tree (Fig. 1a) was built with the available *X. fastidiosa* subsp. *multiplex* genome sequences and using *X. fastidiosa* subsp. *fastidiosa* strains from Mallorca Island, Spain (24), and M23 from California, USA, as an outgroup. We note that *X. fastidiosa* subsp. *fastidiosa* strains from Mallorca Island were introduced into the region and currently represent the only report of *X. fastidiosa* subsp. *fastidiosa* in Europe (24). We also performed an ancestral state reconstruction using strain geographic location as a character. The results indicate that the ancestral nodes within *X. fastidiosa* subsp. *multiplex* originated in the southeastern United States, suggesting that isolates from California, Europe, and Brazil are the consequence of introductions. Although the relationships presented in the phylogenetic tree were supported by high bootstrap values, the directionality of pathogen dissemination is complex. For instance, the introduction of *X. fastidiosa* subsp. *multiplex* into Brazil (RAAR14 plum 327) from the United States was first reported in the 1970s, which constitutes the only reported case of *X. fastidiosa* subsp. *multiplex* in South America (25). Similarly, isolates from Tuscany, Italy, might have been introduced from California or from the southeastern United States (Figure 1a). We note that strain polygala was collected from an urban garden in California not geographically near agricultural areas in the state and may have itself been an introduction from another region (i.e., *Polygala myrtifolia* is an ornamental plant that has been associated with most *X. fastidiosa* outbreaks in Europe, regardless of location and strain phylogenetic placement).

Another clade of *X. fastidiosa* subsp. *multiplex* (including ST6, ST7, and ST81) associated with introductions into Europe has, in fact, evidence of four independent introduction events into three distinct regions (the island of Corsica, the Menorca and Mallorca islands, and Alicante, in mainland Spain). This monophyletic clade has limited diversity compared to *X. fastidiosa* subsp. *multiplex* at large. Furthermore, based on the likelihood of the ancestral state reconstruction, we propose that this clade is also associated with introductions into California, USA, Europe, and Georgia, USA (strain Griffin1 [26], was isolated from an oak tree, *Quercus rubra*). The interpretation that this clade is endemic to Europe and expanded via multiple introductions into the United States and within Europe, while possible, is unlikely. A parsimonious interpretation of the phylogeny indicates that *X. fastidiosa* subsp. *multiplex* was introduced into the

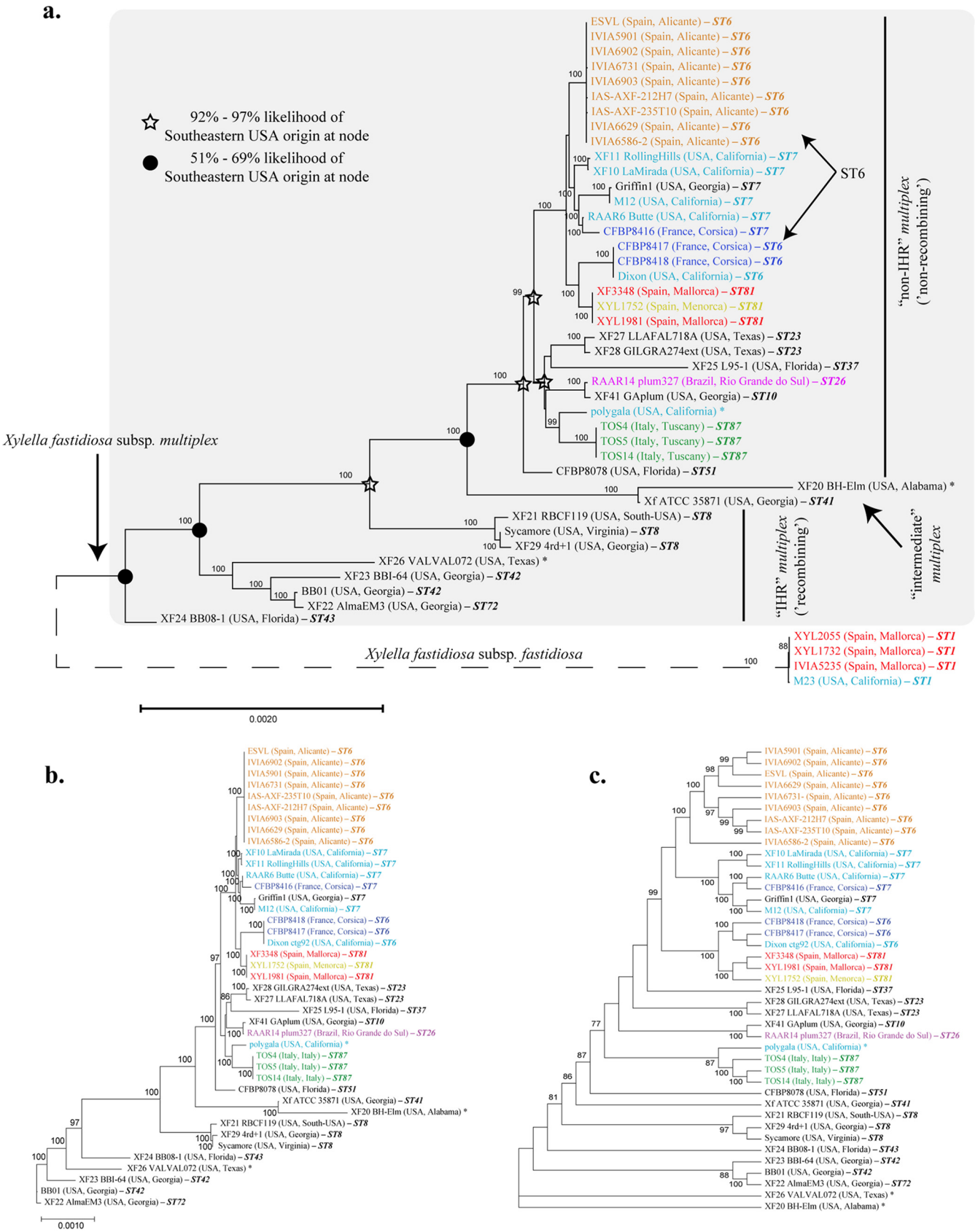

FIG 1 Phylogenetic trees and cladograms showing the complex relationship between isolates from the American and European continents. Font colors correspond to geographic location: southeastern United States (black), California, USA (light blue), Brazil (pink), Corsica, France (dark blue), Tuscany, Italy (green),

(Continued on next page)

French island of Corsica twice, once into the Menorca and Mallorca islands in Spain, and one last time into Alicante, in mainland Spain. Additionally, our results showed that there is a lack of ST monophyly for isolates assigned to ST6 using MLST analysis. We found that isolates from Alicante that clustered closer to isolates belonging to ST7, and isolates belonging to the original ST6 from California and France clustered in a paraphyletic group closer to isolates belonging to ST81 from Menorca and Mallorca. These inferences were based on core genomic data including regions subject to recombination. Omitting the recombinant regions from the data set reduced the alignment length by ~24% (250,533 bp with recombinant regions versus 192,126 bp without recombinant regions). The phylogenetic tree generated without recombinant loci was less informative. Nonetheless, when the analyses were repeated only for *X. fastidiosa* subsp. *multiplex* (without *X. fastidiosa* subsp. *fastidiosa*), phylogenetic trees with and without recombination produced the same topological relations, particularly within the "non-IHR" group compared to the "IHR" group (i.e., isolates showing intersubspecific homologous recombination) (Fig. 1b and c).

Evidence of recent introductions of *X. fastidiosa* subsp. *multiplex* into Europe. The single-nucleotide polymorphism (SNP) analysis of 41 *X. fastidiosa* subsp. *multiplex* genomes, including coding sequences and intergenic regions, revealed 5,630 core SNPs and 23,437 noncore SNPs. The ML phylogenetic tree built with the core SNP matrix (Fig. 2) produced the same topology as that obtained from the same 41 core genome sequences (Fig. 1a), with a few exceptions (isolates XF27 and XF28 from the southeastern United States), and most of the internal nodes were supported by high bootstrap values. We found that the clade comprising the ancestral nodes originated in the Southeast United States, and the IHR *X. fastidiosa* subsp. *multiplex* strains and "intermediate-IHR" *X. fastidiosa* subsp. *multiplex* strains are supported by 190 common core SNPs, whereas the clade comprising the non-IHR *X. fastidiosa* subsp. *multiplex* strains share 117 core SNPs. Strains from Alicante, Tuscany, and the Menorca and Mallorca islands form compact clusters sharing, respectively, 79, 129, and 41 core SNPs. The overall genetic diversity of isolates from Alicante, the Menorca and Mallorca islands, and Tuscany was very low, as inferred by the short branch lengths and the very low number of differentiating SNPs within each subpopulation (10 SNPs for strains from Alicante, 3 SNPs for strains from the Menorca and Mallorca islands, and only 1 SNP for the Tuscany strains). These findings are indicative of a recent introduction in each of those areas. These results were further supported by nucleotide diversity data which indicate that the isolates within the southeastern United States are far more genetically diverse (nucleotide diversity [π] = 354.40) than isolates from California (π = 59.53) or Europe as a whole (π = 54.27). Results also showed that the nucleotide diversity between California, USA, and Europe is low (π = 60.06), and each of those regions/populations show similar genetic diversities compared to the Southeast United States (southeastern United States versus California, π = 289.21; southeastern United States versus Europe, π = 293.81). While the smaller sample size in the California population could bias nucleotide diversity estimates, the sample size between the European and southeastern populations is similar. Thus, the reported differences in nucleotide diversity are likely the product of biological and evolutionary processes and not the result of sample size biases.

FIG 1 Legend (Continued)
Mallorca, Spain (red), Menorca, Spain (yellow), and Alicante, Spain (orange). The sequence type (ST) is included on each isolate name. Isolates without a sequence type are marked with an asterisk (*). Bootstrap values indicate node confidence, with values lower than 70 not shown. (a) An ML phylogeny built using a core genome alignment (311 genes, 250,533 bp) from 41 draft or finished whole-genome sequences shows the evolutionary relationship between isolates from the European and American continents. *X. fastidiosa* subsp. *fastidiosa* isolates are used as an outgroup. *X. fastidiosa* subsp. *multiplex* isolates (shaded in gray) show a complex relationship, suggestive of multiple introductions to the European continent from the Americas. *X. fastidiosa* subsp. *multiplex* isolates are divided into two main groups, the IHR group (isolates showing intersubspecific homologous recombination) at the base of the tree (formed by most southeastern U.S. isolates) and the recently divergent and non-IHR group (48). Isolate XF20 BH-Elm is described as an intermediate recombinant. The likelihood of the southeastern United States as the point of origin was obtained via an ancestral reconstruction analysis. Filled circles, 51 to 69% likelihood; stars, 92 to 97% likelihood. (b) An ML unrooted phylogeny built using a core genome alignment (348 genes, 272,908 bp) shows the evolutionary relationship between *X. fastidiosa* subsp. *multiplex* isolates. Node support and tree topology are largely identical to those seen in panel A when *X. fastidiosa* subsp. *fastidiosa* is included. (c) ML unrooted cladogram built using a core genome alignment after removal of recombinant segments (180,549 bp). The cladogram was used to facilitate tree display due to short branch lengths. Tree topology is largely conserved within the non-IHR group; multiple changes are observed within the IHR group.

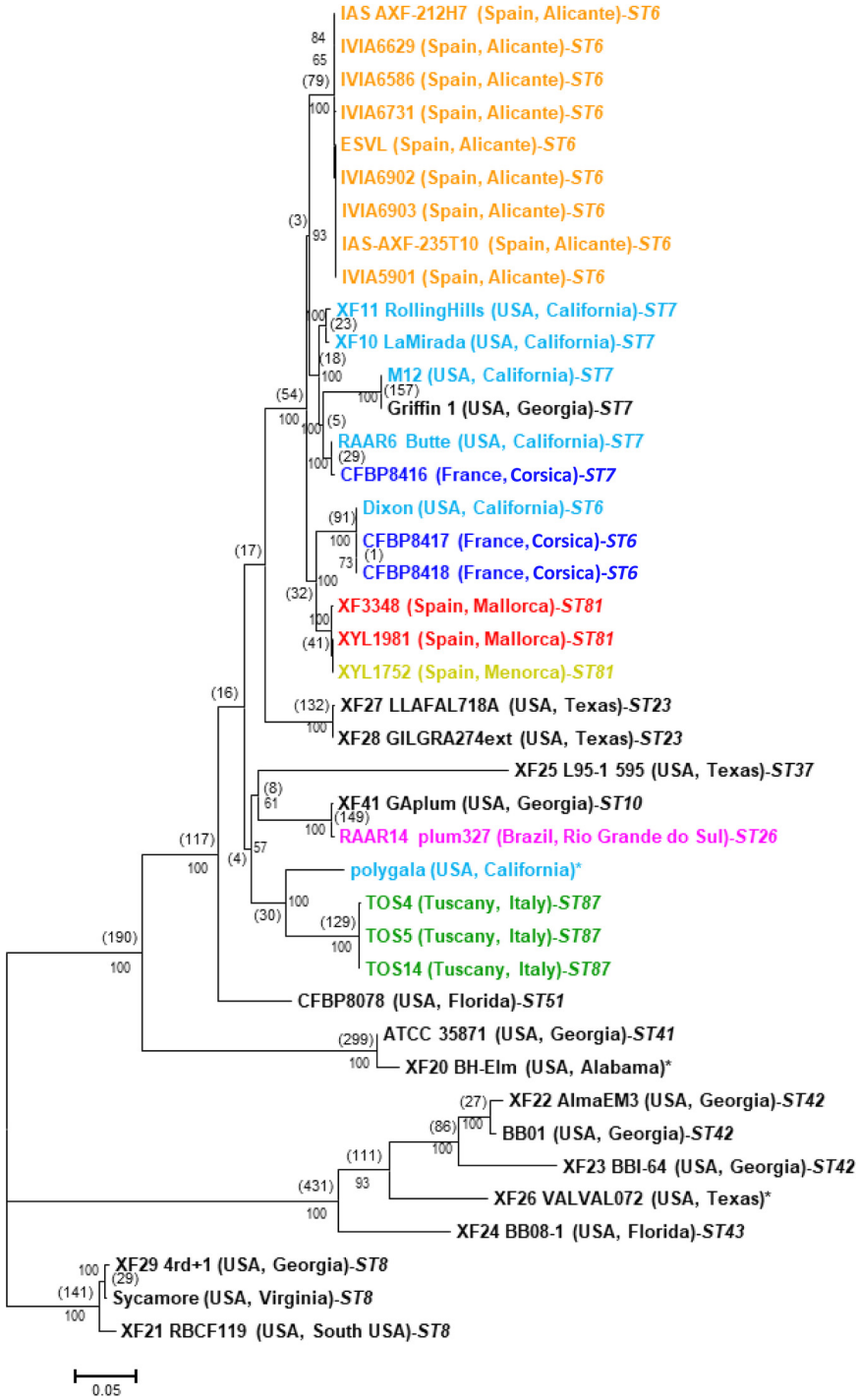

FIG 2 ML phylogeny showing the genetic relationships among *X. fastidiosa* subsp. *multiplex* isolates reconstructed using 5,630 core single-nucleotide polymorphisms (SNPs) from 41 draft or finished whole-genome sequences. The numbers on the branches indicate the values of the bootstrap analyses. The numbers between brackets in each node indicate the numbers of SNPs exclusive of the cluster. Font colors correspond to geographic location: Corsica, France (dark blue), Tuscany, Italy (green), Mallorca, Spain (red), Menorca, Spain (yellow), and Alicante, Spain (orange).

Recombination detected among strains may predate presence in Europe. A recombination detection analysis was performed among European isolates of *X. fastidiosa* subsp. *fastidiosa* and subsp. *multiplex*. A total of 104 *X. fastidiosa* subsp. *fastidiosa* genes and 74 *X. fastidiosa* subsp. *multiplex* genes were located within recombinant

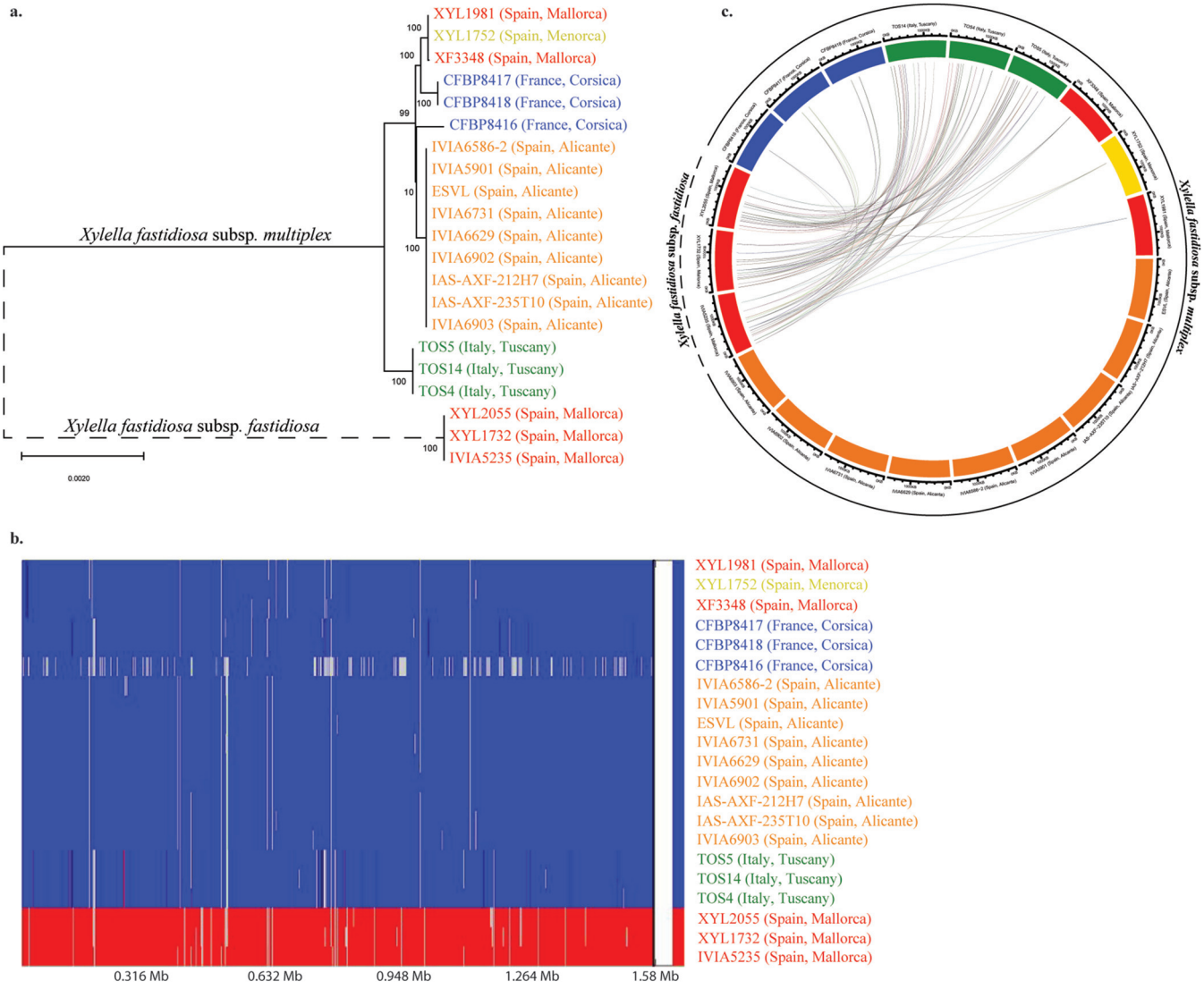

FIG 3 Phylogenetic trees and plots show recombination patterns between *Xylella fastidiosa* subsp. *fastidiosa* and subsp. *multiplex* isolates. Font colors correspond to geographic location: Corsica, France (dark blue), Tuscany, Italy (green), Mallorca, Spain (red), Menorca, Spain (yellow), and Alicante, Spain (orange). (a) ML phylogeny built using a core genome alignment (1,588 genes, 1,569,508 bp); *X. fastidiosa* subsp. *fastidiosa* isolates are used as an outgroup. Bootstrap values indicate node confidence, with values lower than 70 not being shown. (b) FastGEAR plot output showing recent or strain-specific recombination events between European isolates. The plot shows two evolutionarily distinct groups, *X. fastidiosa* subsp. *fastidiosa* (red) and *X. fastidiosa* subsp. *multiplex* (blue). Colored bars show donor and recipient recombinant segments along the length of the core genome alignment. White spaces indicate gaps. (c) Circle plot showing the frequency of recombination events and their location along the length of the core genome alignment of European isolates. Each line indicates a recombinant segment, with the thickness of the line indicating its length.

regions (see Table S1 in the supplemental material). Most of the recombinant genes identified were classified as hypothetical proteins with unknown functions. Several recent (or strain-specific) recombination events were detected among European strains, with *X. fastidiosa* subsp. *fastidiosa* sequences acting as donors to individual *X. fastidiosa* subsp. *multiplex* organisms (Fig. 3b and c). No ancestral (or lineage-specific) recombination events were detected. Interestingly, no recombinant events were detected with sequences from Alicante. Among the observed recombinant regions, *X. fastidiosa* subsp. *fastidiosa* sequences from Mallorca Island, Spain, and *X. fastidiosa* subsp. *multiplex* sequences from Tuscany, Italy, shared the highest number of events. Fewer events were detected between *X. fastidiosa* subsp. *multiplex* sequences from Corsica and *X. fastidiosa* subsp. *fastidiosa* (Fig. 3c). Few events were observed between *X. fastidiosa* subsp. *fastidiosa* and *X. fastidiosa* subsp. *multiplex* strains from Mallorca and Menorca in Spain; Mallorca Island is the only place in Europe where the two subspecies are known

to be sympatric. While the algorithm identified these recombination events as "recent," it is worth noting that there was more evidence of recombination between allopatric populations of different *X. fastidiosa* subspecies (the Menorca and Mallorca islands in Spain and Tuscany in Italy) than sympatric populations (in Mallorca Island). It is possible that that is the case, but we interpret these recent events as exchanges that occurred in North America prior to the introduction of these clades into Europe.

DISCUSSION

Pathogen pollution is a leading driver of emerging plant diseases (6). We used whole-genome sequence data to analyze long-distance spread of the plant pathogen *X. fastidiosa*, focusing on *X. fastidiosa* subsp. *multiplex*. This analysis also included strains obtained from natural environments in Europe. Based on available data, results demonstrate several introductions of *X. fastidiosa* subsp. *multiplex* into Europe and one into Brazil, all with origins in the United States. The introduction into Brazil may, in fact, be associated with the spread of an earlier introduction into South America, although details remain unclear due to the lack of data (25). We interpret these results as a consequence of commercial trade of plant material infected with *X. fastidiosa*. Because *X. fastidiosa*, as a species, infects hundreds of plants, primarily without disease symptoms (12, 13), and insect vectors are unlikely to disseminate at continental scales, it is reasonable to assume that anthropogenic activities are linked to the introductions of these pathogen strains into Europe (e.g., 1, 12).

While it was expected that whole-genome sequence data would assist in the phylogenetic resolution of *X. fastidiosa* strains and related clades (27, 28), it has been generally accepted that MLST is sufficient to discern relevant biological groups and dispersal pathways (17). We show that this may not be always the case. First, we demonstrate that STs may not be monophyletic, illustrated here by the paraphyletic clade in which all ST6 strains were included. MLST is useful in resolving taxonomic discrepancies for bacterial species with high recombination rates, such as *X. fastidiosa* (18); however, issues related to phylogenetic resolution have been observed in the past with MLST, leading to proposals such as the "fuzzy species" concept (29). This has a number of implications in regard to the use of MLST for phylogenetic purposes but is also of applied relevance for the identification of potential host plant species susceptible to particular STs. Lastly, studies have suggested that there is no congruence between *X. fastidiosa* genotype and host plant species phylogenies (7, 21). Our results concerning the lack of ST monophyly raise additional questions about the phylogenetic resolution appropriate for *X. fastidiosa* eradication, quarantine, and trade-related decision-making.

The use of strains belonging to *X. fastidiosa* subsp. *fastidiosa* as an outgroup, notably isolates from grapevine in the United States, and those recovered from grapevine in the Balearic Island of Mallorca (a recent introduction as well), allowed for insights into the evolution and diversity of *X. fastidiosa* subsp. *multiplex*. Nunney et al. (23) demonstrated that there are two major groups within *X. fastidiosa* subsp. *multiplex*, defined as one with high recombination rates ("intersubspecific recombination" or IHR clade), another with limited evidence of recombination (non-IHR clade), and an "intermediate" clade between IHR and non-IHR. The non-IHR *X. fastidiosa* subsp. *multiplex* nonrecombining group is a monophyletic clade, while the IHR *X. fastidiosa* subsp. *multiplex* group is not a clade but is paraphyletic, representing three separate *X. fastidiosa* subsp. *multiplex* clades stemming from the three most ancestral nodes of the *X. fastidiosa* subsp. *multiplex* phylogeny. Vanhove et al. (18) recovered a similar pattern at the genome scale, also demonstrating that the recombining group of *X. fastidiosa* subsp. *multiplex* was the recipient of several DNA fragments from *X. fastidiosa* subsp. *fastidiosa*. The rate of substitution accumulation in our data set, i.e., branch lengths, as well as the paraphyly in the IHR clade, suggest an alternative explanation for the two groups previously proposed. It is possible that instead of one group being more subject to recombination than another, the non-IHR clade (to which all nonsoutheastern U.S. strains belong) is much younger, as represented by the lower levels of genetic diversity.

In other words, it is possible that detectable recombination events are less frequent in these strains because they belong to a more recent lineage. While this hypothesis was not explicitly tested, a phylogenetic tree that did not include recombining regions also showed that branch lengths for the IHR clade were substantially longer than for the non-IHR clade (Fig. 1c). Interestingly, the geography-based ancestral state reconstruction analysis performed suggested that *X. fastidiosa* subsp. *multiplex* originates from the southeastern United States and that it was more recently spread into western states such as California. Since ancestral/endemic populations tend to be more genetically diverse than introduced ones, and the nucleotide diversity of both California, USA, and Europe is similar, these results also support the notion that both populations are relatively young. Moreover, nucleotide diversity estimates also varied within Europe. Except for the island of Corsica ($\pi = 36$), all nucleotide diversity estimates ranged from 1e-05 (Tuscany, in Italy, and the Menorca and Mallorca islands, in Spain) to 0.2 (Alicante, in mainland Spain). This pattern supports the hypotheses of recent multiple introductions to Corsica compared to the single recent introductions to Italy and Spain. Due to the lack of older isolates, the data set used in this study did not allow the implementation of a tip-dating-based molecular clock; nonetheless, it is possible that additional data will assist in testing this hypothesis. It is worth noting that all strains introduced into Europe are from the non-IHR clade, although it is not evident what that may indicate.

Although *X. fastidiosa* subsp. *multiplex* has been successfully established in several nonnative regions, patterns of spread cannot be easily inferred. We assume that *X. fastidiosa* subsp. *multiplex* originates from the southeastern United States, following likelihoods derived from the ancestral state reconstruction, overall branch lengths, and basal placement of southeastern U.S. strains within the clade (i.e., subspecies). As such, introductions occurred into Brazil (RAAR14 plum327), California (polygala), and Tuscany, Italy (TOS strains). However, it is not possible to determine if the introduction into Tuscany originated in California or if the polygala strain itself was introduced into California independently. The latter strain was obtained from an ornamental plant in an urban yard on a host plant species considered an indicator species for *X. fastidiosa* surveillance in Europe (12). In other words, it is possible that this clade lacks representatives from the area of origin. The other clade includes strains from two independent introductions into Spain (Alicante and the Mallorca/Menorca islands), another two into the island of Corsica, and isolates from California and Georgia, USA. While the multiple introductions into Europe are evident, dispersal paths are not always clear. In one case, strain Griffin1 was obtained from oak in the southeastern United States, potentially a natural host of the strain in its original range. Based on the ancestral state reconstruction, there is a 99% likelihood that Griffin1 was introduced from California back into the southeastern United States. M12, its sister taxon, was obtained from almond commercially grown in California. On the other hand, between Europe and the United States, there are difficulties in interpreting the directionality of introduction at other nodes within this clade. Generally, considering the likelihood of strains being introduced and established at one location and then being reintroduced back to the region of origin, the most parsimonious interpretation of the data is that *X. fastidiosa* subsp. *multiplex* was introduced into California, where it spread in almond and other crop and noncrop plant species, and was then introduced into France and Spain via the commercial trade of plants.

Soubeyrand et al. (30) proposed two scenarios for the emergence of *X. fastidiosa* subsp. *multiplex* in Corsica based on mathematical models, the first with an introduction in ~2001 (1998 to 2005) and the second with an introduction in ~1985 (1978 to 1993), also supporting the hypothesis that this is a young clade. The lack of a robust molecular clock for this data set does not permit inferences on when introductions occurred, but the lack of diversity within the Alicante, the Mallorca and Menorca, and the Tuscany populations indicates that these are relatively recent introductions. It should be noted that trade of plants such as almond trees is currently highly regulated, so these introductions either occurred prior to enforcement of regulations, through

other nonsusceptible plant species, or through illegal introductions. Unfortunately, *X. fastidiosa* was considered to be a virus until the 1970s and was poorly studied until the late 1990s (31), thus impacting efforts to limit its spread to new geographic regions.

Recombination events may be recent (i.e., between strains in a population) or ancestral (i.e., events that are fixed in a population). Ample evidence of recombination in *X. fastidiosa* populations has led to suggestions that the process may be associated with the emergence of novel diseases (e.g., 15, 23, 25, 32). Therefore, there is substantial concern that the continuous introductions of *X. fastidiosa* into new geographic regions present risks beyond the inherent historical threats associated with the genotypes introduced (17, 32). Using genome sequences for strains present in Europe, including both *X. fastidiosa* subsp. *multiplex* and *X. fastidiosa* subsp. *fastidiosa*, we attempted to determine if there was evidence of recent recombination among those populations. While recombination was detected in the data set, it is not clear if those events occurred prior to or after pathogen establishment in Europe, since all events involved strains from *X. fastidiosa* subsp. *fastidiosa* in the Menorca and Mallorca islands. There was one shared recombination event with *X. fastidiosa* subsp. *multiplex* in that archipelago, which is present in all isolates. Similarly, all other events are shared among all strains within a particular population, suggesting that these are all ancestral and occurred prior to these introductions. It is worth noting that the analyses indicate that strain CFPB8416 (ST7) has recombination events with a genotype(s) not included in the analyses, further suggesting its distinction from the other strains from Corsica.

In summary, our results indicate that *X. fastidiosa* subsp. *multiplex* has been introduced multiple times into Europe and once into Brazil. It appears that most introductions into Europe originated from California, although *X. fastidiosa* subsp. *multiplex* itself was introduced into California, and the phylogenetic placement of strain polygala suggests that novel introductions continue to occur into the western United States. There is evidence of recombination between *X. fastidiosa* sympatric populations (e.g., 27, 33), but there is no conclusive evidence that it has already occurred in Europe. In addition, the data show that MLST-based STs are not monophyletic. This study demonstrates the need to utilize whole-genome sequences to study pathogen introductions at outbreak stages, where a limited number of genetic markers do not provide sufficient phylogenetic resolution to determine paths of dispersal or relationships among strains that are of biological and quarantine relevance. Overall, the work illustrates the risks associated with the commercial trade of plant material at global scales and the need to develop effective policy to limit the likelihood of pathogen pollution into naive regions.

MATERIALS AND METHODS

Bacterial isolate data and sequencing. Table 1 contains metadata for all of the published and unpublished isolates included in this study. New *X. fastidiosa* subsp. *multiplex* isolates were sequenced with the Illumina HiSeq 4000 platform. The raw reads and metadata for newly sequenced isolates are publicly available at the GenBank SRA database (Table 1). Sequence quality was assessed using MultiQC (34). Low-quality reads and adapter sequences were removed from all raw reads using seqtk v1.2 (35) and Cutadapt v1.14 (36), respectively. After preprocessing, the genomes of the isolates sequenced were assembled *de novo* with SPAdes v3.13 (37, 38). Assembled contigs were reordered using the Mauve contig mover function (39). Complete publicly available assemblies were used as references; specifically, *X. fastidiosa* subsp. *fastidiosa* scaffolds were reordered using the Temecula 1 assembly (GCA_000007245), while *X. fastidiosa* subsp. *multiplex* scaffolds were reordered using the M12 assemblies (ASM1932v1; 40) as references, respectively. Assembled and reordered genomes were then annotated using Prokka (41). Depth coverage, coverage, and other assembly-related data for all new isolates are presented in Table 2.

Core genome analyses. Roary v3.11.2 (42) was used to obtain the core, soft-core, shell, and cloud genomes of *X. fastidiosa* subsp. *fastidiosa* and subsp. *multiplex*. A core genome alignment was created using Prokka annotations for isolates from Europe,

TABLE 1 Isolates of *Xylella fastidiosa* used in this study

Isolate name	*X. fastidiosa* subsp.	ST[a]	Host	Geographical origin	Yr of collection	Reference	Accession no. (WGS[b] project)	Accession no. (BioSample)
XYL1732	*fastidiosa*	1	*Vitis vinifera*	Mallorca, Spain	2017	51	QTJT00000000	SAMN09767243
XYL2055	*fastidiosa*	1	*Vitis vinifera*	Mallorca, Spain	2017	51	QTJS00000000	SAMN09767242
IVIA5235	*fastidiosa*	1	*Prunus avium*	Mallorca, Spain	2017	11	QWLC00000000	SAMN09925804
M23	*fastidiosa*	1	*Prunus dulcis*	San Joaquin Valley, CA, USA	2003	40	CP001011	SAMN02598408
CFBP8417	*multiplex*	6	*Spartium junceum*	Alata, Corse-du-Sud, Corsica, France	2015	52	LUYB00000000	SAMN04546482
CFBP8418	*multiplex*	6	*Spartium junceum*	Alata, Corse-du-Sud, Corsica, France	2015	52	LUYA00000000	SAMN04546487
ESVL	*multiplex*	6	*Prunus dulcis*	Benimantell, Spain (Alicante)	2017	10	NZ_QPQV00000000	SAMN09703576
IVIA5901	*multiplex*	6	*Prunus dulcis*	Bolulla, Spain (Alicante)	2017	10	NZ_QPQW00000000	SAMN09703973
IVIA6586-2	*multiplex*	6	*Helichrysum italicum*	Beniardà, Spain (Alicante)	2018	This study	VDCM00000000	SAMN11849688
IAS-AXF-212H7	*multiplex*	6	*Prunus dulcis*	Benimantell, Spain (Alicante)	2018	This study	VCPQ00000000	SAMN11878160
IVIA6629	*multiplex*	6	*Rhamnus alaternus*	Spain, Callosa d'en Sarrià (Alicante)	2018	This study	VCPM00000000	SAMN11866308
IVIA6902	*multiplex*	6	*Prunus dulcis*	Castell de Castells, Spain (Alicante)	2018	This study	VCPO00000000	SAMN11866375
IVIA6903	*multiplex*	6	*Prunus dulcis*	Castell de Castells, Spain (Alicante)	2018	This study	VCPP00000000	SAMN11867883
IAS-AXF-235T10	*multiplex*	6	*Prunus dulcis*	El Castell de Guadalest, Spain (Alicante)	2018	This study	VCQO00000000	SAMN11881272
IVIA6731	*multiplex*	6	*Helichrysum italicum*	Tárbena, Spain (Alicante)	2018	This study	VCPN00000000	SAMN11866310
Dixon	*multiplex*	6	*Prunus dulcis*	Solano County, CA, USA	2002	53	AAAL00000000	SAMN02441075
RAAR14 plum327	*multiplex*	26	*Prunus domestica*	Veranopolis, Rio Grande do Sul, Brazil	2010	This study	VDDF00000000	SAMN11936157
CFBP8416	*multiplex*	7	*Polygala myrtifolia*	Propriano, Corse-du-Sud, Corsica, France	2015	52	LUYC00000000	SAMN04546448
RAAR6 Butte	*multiplex*	7	*Prunus dulcis*	Butte County, CA, USA	1990s	This study	VDDE00000000	SAMN11936022
Xf10 LaMirada	*multiplex*	7	*Olea europaea*	La Mirada, CA, USA	2011	18		SAMN09941418
Xf11 RollingHills	*multiplex*	7	*Olea europaea*	Rolling Hills, CA, USA	2011	18		SAMN09941419
M12	*multiplex*	7	*Prunus dulcis*	San Joaquin Valley, Kern County, CA, USA	2003	26	CP000941	SAMN02598402
Griffin1	*multiplex*	7	*Quercus rubra*	Griffin, GA, USA	2006	26	AVGA00000000	SAMN02472062
Xf29 4rd+1	*multiplex*	8	*Platanus* sp.	Athens, Clarke County, GA, USA	Unknown	18		SAMN09941429
Xf21 RBCF119	*multiplex*	8	*Carya illinoinensis*	Southeast USA	2012	18		SAMN09941421
Sycamore	*multiplex*	8	*Platanus occidentalis*	Virginia, USA	2002	26	JMHP00000000	SAMN02709772
Xf41 GAplum	*multiplex*	10	*Prunus domestica*	Georgia, USA	Unknown	18		SAMN09941431
Xf28 GILGRA274ext	*multiplex*	23	*Helianthus* sp.	Texas, USA	Unknown	18		SAMN09941428
Xf27 LLAFAL718A	*multiplex*	23	*Iva annua*	Texas, USA	Unknown	18		SAMN09941427
Xf25 L95-1	*multiplex*	37	*Lupinus aridorum*	Florida, USA	1995	18		SAMN09941425
Xf ATCC 35871	*multiplex*	41	*Prunus salicina*	Georgia, USA	1983		AUAJ00000000	SAMN02441559
BB01	*multiplex*	42	*Vaccinium corymbosum*	Nahunta, GA, USA	2016	54	MPAZ00000000	SAMN05982167
Xf23 BBI-64	*multiplex*	42	*Vaccinium* sp.	Hoboken, Brantley County, GA, USA	2009	18	PUJA00000000	SAMN08537147
Xf22 AlmaEM3	*multiplex*	42	*Vaccinium* sp.	Alma, Bacon County, GA, USA	2011	18	PUIY00000000	SAMN08537149
Xf24 BB08-1	*multiplex*	43	*Vaccinium* sp.	Palatka, Putnam County, FL, USA	2008	18	PUIZ00000000	SAMN08537148
CFBP8078	*multiplex*	51	*Vinca* sp.	Florida, USA	1987	55	PHFS00000000	SAMN07999164
XF3348	*multiplex*	81	*Prunus dulcis*	Mallorca, Spain	2018	This study	VDCL00000000	SAMN11891224
XYL1981	*multiplex*	81	*Ficus carica*	Mallorca, Spain	2018	This study	VDCJ00000000	SAMN11891201
XYL1752	*multiplex*	81	*Prunus dulcis*	Menorca, Spain	2018	This study	VDCK00000000	SAMN11891208
TOS4	*multiplex*	87	*Prunus dulcis*	Tuscany, Italy	2018	9	SMTH00000000	SAMN11104854
TOS5	*multiplex*	87	*Polygala myrtifolia*	Tuscany, Italy	2018	9	SMTI00000000	SAMN11104883
TOS14	*multiplex*	87	*Spartium junceum*	Tuscany, Italy	2018	9	SMTJ00000000	SAMN11104922
Xf26 VALVAL072	*multiplex*		*Ambrosia sp.*	USA (likely Texas)	Unknown	18		SAMN09941426
Xf20 BH-Elm	*multiplex*		*Ulmus americana*	Birmingham, AL, USA	2010	18		SAMN09941420
Polygala	*multiplex*		*Polygala myrtifolia*	Oakland, CA, USA	2016	18		SAMN09941407

[a]Sequence type (ST) obtained from the *Xylella fastidiosa* MLST database (https://pubmlst.org/xfastidiosa/).
[b]WGS, whole-genome sequencing.

TABLE 2 Assembly information from the strains sequenced in this study

Isolate name	No. of contigs	Total length (bp)	Genome coverage (×)	N_{50} (bp)	L_{50} (kpb)	Maximum contig length (bp)	GC content (%)	No. of predicted genes (Prokka)
IVIA6586-2	341	2,578,155	941	78,73	9	251.315	51.8	2,431
IAS-AXF-212H7	336	2,522,542	932	103,609	7	309.947	51.8	2,376
IVIA6629	1,099	2,731,703	890	96,774	8	309.951	53.2	2,429
IVIA6902	1,584	2,870,108	1177	94,781	9	309.939	53.5	2,510
IVIA6903	534	2,602,829	787	103,089	7	309.946	52.1	2,410
IAS-AXF-235T10	432	2,550,209	1159	103,901	7	309.939	51.7	2,399
IVIA6731	422	2,598,893	782	98,660	7	309.939	51.9	2,441
RAAR14 plum327	192	2,543,559	776	96,594	8	342.38	51.6	2,471
RAAR6 Butte	140	2,466,226	719	128.621	7	311.304	51.8	2,337
XF3348	392	2,573,395	914	102,161	7	257.114	51.8	2,422
XYL1981	353	2,554,510	757	96,968	9	275.285	51.8	2,409
XYL1752	422	2,579,375	895	98,273	9	245.39	51.8	2,429

Brazil, and the United States (Table 1 and 2). Subsequently, a maximum likelihood (ML) tree was built with RAxML (43) using the GTRCAT substitution model. Tree topology and branch support were assessed using 1,000 bootstrap replicates. Similarly, a core genome alignment and ML tree were built only for isolates of *X. fastidiosa* subsp. *multiplex*. In addition, recombinant regions were later identified and removed from the *X. fastidiosa* subsp. *multiplex* core alignment and used to build a third ML cladogram. A final core gene alignment and ML tree were built for *X. fastidiosa* subsp. *fastidiosa* and subsp. *multiplex* isolates obtained from European infected plants.

The pan-genome of *X. fastidiosa* subsp. *fastidiosa* and subsp. *multiplex* isolates was inferred with Roary v3.11.2 (42). Roary was used to produce a gene presence/absence matrix that was visualized in Phandango (www.phandango.net), and an accessory tree based on the binary accessory presence/absence matrix was built and plotted using the binary model at IQ-TREE (http://iqtree.cibiv.univie.ac.at/). The branches of the tree were proportionally transformed with Figtree (http://tree.bio.ed.ac.uk/software/figtree/).

Single-nucleotide polymorphism analysis. Single-nucleotide polymorphisms (SNPs) were identified from a total of 41 *X. fastidiosa* subsp. *multiplex* genomes using kSNP3.02 software (44) with annotated nucleotide GenBank files as input data. SNP discovery was based on k-mer analysis (21_kmer), i.e., single variant positions within sequences of nucleotide length k. A core SNP matrix was generated and used as input in MEGA7 alignment software. This matrix was subsequently analyzed using the maximum likelihood (ML) method. Robustness of the ML tree was assessed using 1,000 bootstrap replications. Additionally, nucleotide diversity (π) within and between populations defined based on geographical regions (southeastern United States, California, and Europe) estimated from the core genome alignment was calculated using the PopGenome v2.7.1 package in R (45).

Phylogenetic analyses and geographic ancestral state reconstruction. Six geographic character states were coded as possible locations for the ancestor represented at each internal node of the phylogeny. These six regions were Spain, Italy, France, Brazil, southeastern United States, and California, USA. The marginal ancestral state likelihood estimates of each location for all internal nodes of the ML phylogenetic tree were calculated using the rerooting method of Yang et al. (46) in the R package Phytools (47) and mapped using the package APE (48). The ML estimates at each node were calculated based on both the equal rate transition model (i.e., fixed rate of geographic change between any locations) and the symmetrical rate transition model (i.e., fixed rates of geographic change symmetrically pairwise between locations but not between all locations). The fit of the two models to the data was compared using the Akaike information criterion (AIC). The symmetrical rates model had a lower AIC value than the equal rates model; thus, it was chosen for ML estimates at internal nodes.

Recombination detection. Ancestral and recent recombinant regions in the core genome alignment were identified using fastGEAR with default parameters (49). Recombination levels were detected between lineages (ancestral recombination) and

between strains (recent recombination). Donor and recipient recombinant regions were visualized using the fastGEAR plotRecombinations script. Additionally, recent recombinant regions were mapped using the R package Circlize (50). An in-house Python v3.6 script was used to identify genes contained entirely within identified recombinant segments. Recombinant genes were identified using the publicly available genomes XYL1732 (51, 24) for *X. fastidiosa* subsp. *fastidiosa* and CFBP8416 (52) for *X. fastidiosa* subsp. *multiplex* (Table 1).

Data availability. Newly determined data accession numbers for IVIA6586-2, IAS-AXF-212H7, IVIA6629, IVIA6902, IVIA6903, IAS-AXF-235T10, IVIA6731, RAAR14 plum327, RAAR6 Butte, XF3348, XYL1981, and XYL1752 are shown in Table 1.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, XLSX file, 0.02 MB.

ACKNOWLEDGMENTS

We thank members of the IAS-CSIC research group "Phytopathology in Sustainable Agricultural Systems," Adela Monterde and Inmaculada Navarro from IVIA, and Marina Montesinos from SEMILLA, Mallorca, Spain, for technical support.

This work was funded by the following awards: XF-ACTORS (*Xylella fastidiosa* Active Containment Through a Multidisciplinary-Oriented Research Strategy, grant 727987 from the European Union's Horizon 2020 Framework Research Programme), E-RTA2017-00004-C02 (Desarrollo de estrategias de erradicación, contención y control de *X. fastidiosa* en España from "Programa Estatal de I+D+I Orientada a los Retos de la Sociedad of the Spanish Government" and FEDER), the Spanish Olive Oil Interprofessional, and the California Department of Food and Agriculture Pierce' Disease Research Program.

This work used the Vincent J. Coates Genomics Sequencing Laboratory at the University of California, Berkeley, supported by NIH S10 OD018174 instrumentation grant.

REFERENCES

1. Gottwald T, Luo W, Posny D, Riley T, Louws F. 2019. A probabilistic census-travel model to predict introduction sites of exotic plant, animal and human pathogens. Philos Trans R Soc B 374:20180260. https://doi.org/10.1098/rstb.2018.0260.
2. Dudas G, Carvalho LM, Bedford T, Tatem AJ, Baele G, Faria NR, Park DJ, Ladner JT, Arias A, Asogun D, Bielejec F, Caddy SL, Cotten M, D'Ambrozio J, Dellicour S, Di Caro A, Diclaro JW, Duraffour S, Elmore MJ, Fakoli LS, Faye O, Gilbert ML, Gevao SM, Gire S, Gladden-Young A, Gnirke A, Goba A, Grant DS, Haagmans BL, Hiscox JA, Jah U, Kugelman JR, Liu D, Lu J, Malboeuf CM, Mate S, Matthews DA, Matranga CB, Meredith LW, Qu J, Quick J, Pas SD, Phan MVT, Pollakis G, Reusken CB, Sanchez-Lockhart M, Schaffner SF, Schieffelin JS, Sealfon RS, Simon-Loriere E, et al. 2017. Virus genomes reveal factors that spread and sustained the Ebola epidemic. Nature 544:309–315. https://doi.org/10.1038/nature22040.
3. O'Hanlon SJ, Rieux A, Farrer RA, Rosa GM, Waldman B, Bataille A, Kosch TA, Murray KA, Brankovics B, Fumagalli M, Martin MD, Wales N, Alvarado-Rybak M, Bates KA, Berger L, Böll S, Brookes L, Clare F, Courtois EA, Cunningham AA, Doherty-Bone TM, Ghosh P, Gower DJ, Hintz WE, Höglund J, Jenkinson TS, Lin C-F, Laurila A, Loyau A, Martel A, Meurling S, Miaud C, Minting P, Pasmans F, Schmeller DS, Schmidt BR, Shelton JMG, Skerratt LF, Smith F, Soto-Azat C, Spagnoletti M, Tessa G, Toledo LF, Valenzuela-Sánchez A, Verster R, Vörös J, Webb RJ, Wierzbicki C, Wombwell E, Zamudio KR, et al. 2018. Recent Asian origin of chytrid fungi causing global amphibian declines. Science 360:621–627. https://doi.org/10.1126/science.aar1965.
4. Scheele BC, Pasmans F, Skerratt LF, Berger L, Martel A, Beukema W, Acevedo AA, Burrowes PA, Carvalho T, Catenazzi A, De la Riva I, Fisher MC, Flechas SV, Foster CN, Frías-Álvarez P, Garner TWJ, Gratwicke B, Guayasamin JM, Hirschfeld M, Kolby JE, Kosch TA, La Marca E, Lindenmayer DB, Lips KR, Longo AV, Maneyro R, McDonald CA, Mendelson J, Palacios-Rodriguez P, Parra-Olea G, Richards-Zawacki CL, Rödel M-O, Rovito SM, Soto-Azat C, Toledo LF, Voyles J, Weldon C, Whitfield SM, Wilkinson M, Zamudio KR, Canessa S. 2019. Amphibian fungal panzootic causes catastrophic and ongoing loss of biodiversity. Science 363:1459–1463. https://doi.org/10.1126/science.aav0379.
5. Daszak P, Cunningham AA, Hyatt AD. 2000. Emerging infectious diseases of wildlife: threats to biodiversity and human health. Science 287:443–449. https://doi.org/10.1126/science.287.5452.443.
6. Anderson PK, Cunningham AA, Patel NG, Morales FJ, Epstein PR, Daszak P. 2004. Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. Trends Ecol Evol 19:535–544. https://doi.org/10.1016/j.tree.2004.07.021.
7. Sicard A, Zeilinger AR, Vanhove M, Schartel TE, Beal DJ, Daugherty MP, Almeida RPP. 2018. *Xylella fastidiosa*: insights into an emerging plant pathogen. Annu Rev Phytopathol 56:181–202. https://doi.org/10.1146/annurev-phyto-080417-045849.
8. Loconsole G, Saponari M, Boscia D, D'Attoma G, Morelli M, Martelli GP, Almeida RPP. 2016. Intercepted isolates of *Xylella fastidiosa* in Europe reveal novel genetic diversity. Eur J Plant Pathol 146:85–94. https://doi.org/10.1007/s10658-016-0894-x.
9. Giampetruzzi A, D'Attoma G, Zicca S, Abou Kubaa R, Rizzo D, Boscia D, Saldarelli P, Saponari M. 2019. Draft genome sequence resources of three strains (TOS4, TOS5, and TOS14) of *Xylella fastidiosa* infecting different host plants in the newly discovered outbreak in Tuscany, Italy. Phytopathology 109:1516. https://doi.org/10.1094/PHYTO-04-19-0108-A.
10. Giampetruzzi A, Velasco-Amo MP, Marco-Noales E, Montes-Borrego M, Román-Écija M, Navarro I, Monterde A, Barbé S, Almeida RPP, Saldarelli P, Saponari M, Montilon V, Savino VN, Boscia D, Landa BB. 2019. Draft genome resources of two strains ("ESVL" and "IVIA5901") of *Xylella fastidiosa* associated with almond leaf scorch disease in Alicante, Spain. Phytopathol 109:219–221. https://doi.org/10.1094/PHYTO-09-18-0328-A.
11. Landa BB, Velasco-Amo MP, Marco-Noales E, Olmo D, López MM, Navarro I, Monterde A, Barbé S, Montes-Borrego M, Román-Écija M, Sapo-

nari M, Giampetruzzi A. 2018. Draft genome sequence of *Xylella fastidiosa* subsp. *fastidiosa* strain IVIA5235, isolated from *Prunus avium* in Mallorca Island, Spain. Microbiol Resour Announc 7:e01222-18. https://doi.org/10.1128/MRA.01222-18.
12. EFSA PLH Panel (EFSA Panel on Plant Health), Jeger M, Caffier D, Candresse T, Chatzivassiliou E, Dehnen-Schmutz K, Gilioli G, Gregoire J-C, Jaques Miret JA, MacLeod A, Navajas Navarro M, Niere B, Parnell S, Potting R, Rafoss T, Rossi V, Urek G, Van Bruggen A, Van der Werf W, West J, Winter S, Almeida R, Bosco D, Jacques M-A, Landa B, Purcell A, Saponari M, Czwienczek E, Delbianco A, Stancanelli G, Bragard C. 2018. Scientific opinion on the updated pest categorisation of Xylella fastidiosa. EFSA J 16:5357.
13. EFSA PLH Panel (EFSA Panel on Plant Health), Bragard C, Dehnen-Schmutz K, Di Serio F, Gonthier P, Jacques M-A, Jaques Miret JA, Justesen AF, MacLeod A, Magnusson CS, Milonas P, Navas-Cortés JA, Potting R, Reignault PL, Thulke H-H, van der Werf W, Civera AV, Yuen J, Zappalà L, Boscia D, Chapman D, Gilioli G, Krugner R, Mastin A, Simonetto A, Spotti Lopes JR, White S, Abrahantes JC, Delbianco A, Maiorano A, Mosbach-Schulz O, Stancanelli G, Guzzo M, Parnell S. 2019. Update of the Scientific Opinion on the risks to plant health posed by *Xylella fastidiosa* in the EU territory. EFSA J 17:5665.
14. Schaad NW, Postnikova E, Lacy G, Fatmi MB, Chang CJ. 2004. *Xylella fastidiosa* subspecies: *X. fastidiosa* subsp. *piercei*, subsp. nov., *X. fastidiosa* subsp. *multiplex* subsp. nov., and *X. fastidiosa* subsp. *pauca* subsp. nov. Syst Appl Microbiol 27:290–300. https://doi.org/10.1078/0723-2020-00263.
15. Nunney L, Schuenzel EL, Scally M, Bromley RE, Stouthamer R. 2014. Large-scale intersubspecific recombination in the plant-pathogenic bacterium *Xylella fastidiosa* is associated with the host shift to mulberry. Appl Environ Microbiol 80:3025–3033. https://doi.org/10.1128/AEM.04112-13.
16. Schuenzel EL, Scally M, Stouthamer R, Nunney L. 2005. A multigene phylogenetic study of clonal diversity and divergence in North American strains of the plant pathogen *Xylella fastidiosa*. Appl Environ Microbiol 71:3832–3839. https://doi.org/10.1128/AEM.71.7.3832-3839.2005.
17. Almeida RPP, Nunney L. 2015. How do plant diseases caused by *Xylella fastidiosa* emerge? Plant Dis 99:1457–1467. https://doi.org/10.1094/PDIS-02-15-0159-FE.
18. Vanhove M, Retchless AC, Sicard A, Rieux A, Coletta-Filho HD, De La Fuente L, Stenger DC, Almeida RPP. 2019. Genomic diversity and recombination among *Xylella fastidiosa* subspecies. Appl Environ Microbiol 85:e02972-18. https://doi.org/10.1128/AEM.02972-18.
19. Scally M, Schuenzel EL, Stouthamer R, Nunney L. 2005. Multilocus sequence type system for the plant pathogen *Xylella fastidiosa* and relative contributions of recombination and point mutation to clonal diversity. Appl Environ Microbiol 71:8491–8499. https://doi.org/10.1128/AEM.71.12.8491-8499.2005.
20. Nunney L, Yuan X, Bromley R, Hartung J, Montero-Astúa M, Moreira L, Ortiz B, Stouthamer R. 2010. Population genomic analysis of a bacterial plant pathogen: novel insight into the origin of Pierce's disease of grapevine in the US. PLoS One 5:e15488. https://doi.org/10.1371/journal.pone.0015488.
21. Nunney L, Azad H, Stouthamer R. 2019. An experimental test of the host-plant range of non-recombinant strains of North American *Xylella fastidiosa* subsp. *multiplex*. Phytopathology 109:294–300. https://doi.org/10.1094/PHYTO-07-18-0252-FI.
22. Kung SH, Almeida RPP. 2011. Natural competence and recombination in the plant pathogen *Xylella fastidiosa*. Appl Environ Microbiol 77:5278–5284. https://doi.org/10.1128/AEM.00730-11.
23. Nunney L, Hopkins DL, Morano LD, Russell SE, Stouthamer R. 2014. Intersubspecific recombination in *Xylella fastidiosa* strains native to the United States: infection of novel hosts associated with an unsuccessful invasion. Appl Environ Microbiol 80:1159–1169. https://doi.org/10.1128/AEM.02920-13.
24. Moralejo E, Borràs D, Gomila M, Montesinos M, Adrover F, Juan A, Nieto A, Olmo D, Seguí G, Landa BB. 2019. Insights into the epidemiology of Pierce's disease in vineyards of Mallorca, Spain. Plant Pathol 68:1458–1471. https://doi.org/10.1111/ppa.13076.
25. Coletta-Filho HD, Francisco CS, Lopes JRS, Muller C, Almeida RPP. 2017. Homologous recombination and *Xylella fastidiosa* host-pathogen associations in South America. Phytopathology 107:305–312. https://doi.org/10.1094/PHYTO-09-16-0321-R.
26. Chen J, Huang H, Chang CJ, Stenger DC. 2013. Draft genome sequence of *Xylella fastidiosa* subsp. *multiplex* strain Griffin-1 from *Quercus rubra* in Georgia. Genome Announc 1:e00756-13. https://doi.org/10.1128/genomeA.00756-13.
27. Nunney L, Vickerman DB, Bromley RE, Russell SA, Hartman JR, Morano LD, Stouthamer R. 2013. Recent evolutionary radiation and host plant specialization in the *Xylella fastidiosa* subspecies native to the United States. Appl Environ Microbiol 79:2189–2200. https://doi.org/10.1128/AEM.03208-12.
28. Nunney L, Yuan X, Bromley RE, Stouthamer R. 2012. Detecting genetic introgression: high levels of intersubspecific recombination found in *Xylella fastidiosa* in Brazil. Appl Environ Microbiol 78:4702–4714. https://doi.org/10.1128/AEM.01126-12.
29. Hanage WP, Fraser C, Spratt BG. 2005. Fuzzy species in recombinogenic bacteria. BMC Biol 3:6. https://doi.org/10.1186/1741-7007-3-6.
30. Soubeyrand S, de Jerphanion P, Martin O, Saussac M, Manceau C, Hendrikx P, Lannou C. 2018. Inferring pathogen dynamics from temporal count data: the emergence of *Xylella fastidiosa* in France is probably not recent. New Phytol 219:824–836. https://doi.org/10.1111/nph.15177.
31. Purcell A. 2013. Paradigms: examples from the bacterium *Xylella fastidiosa*. Annu Rev Phytopathol 51:339–356. https://doi.org/10.1146/annurev-phyto-082712-102325.
32. Potnis N, Kandel PP, Merfa MV, Retchless AC, Parker JK, Stenger DC, Almeida RPP, Bergsma-Vlami M, Westenberg M, Cobine PA, De La Fuente L. 2019. Patterns of inter- and intrasubspecific homologous recombination inform eco-evolutionary dynamics of *Xylella fastidiosa*. ISME J 13:2319. https://doi.org/10.1038/s41396-019-0423-y.
33. Almeida RPP, Nascimento FE, Chau J, Prado SS, Tsai C-W, Lopes SA, Lopes JRS. 2008. Genetic structure and biology of *Xylella fastidiosa* strains causing disease in citrus and coffee in Brazil. Appl Environ Microbiol 74:3690–3701. https://doi.org/10.1128/AEM.02388-07.
34. Ewels P, Magnusson M, Lundin S, Käller M. 2016. Data and text mining MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32:3047–3048. https://doi.org/10.1093/bioinformatics/btw354.
35. Shen W, Le S, Li Y, Hu F. 2016. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. PLoS One 11:e0163962. https://doi.org/10.1371/journal.pone.0163962.
36. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. Embnet J 17:10–12. https://doi.org/10.14806/ej.17.1.200.
37. Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, Prjibelski AD, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R, Clingenpeel SR, Woyke T, McLean JS, Lasken R, Tesler G, Alekseyev MA, Pevzner PA. 2013. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. J Comput Biol 20:714–737. https://doi.org/10.1089/cmb.2013.0084.
38. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.
39. Rissman AI, Mau B, Biehl BS, Darling AE, Glasner JD, Perna NT. 2009. Reordering contigs of draft genomes using the Mauve aligner. Bioinformatics 25:2071–2073. https://doi.org/10.1093/bioinformatics/btp356.
40. Chen J, Xie G, Han S, Chertkov O, Sims D, Civerolo EL. 2010. Whole genome sequences of two *Xylella fastidiosa* strains (M12 and M23) causing almond leaf scorch disease in California. J Bacteriol 192:4534. https://doi.org/10.1128/JB.00651-10.
41. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
42. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691–3693. https://doi.org/10.1093/bioinformatics/btv421.
43. Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690. https://doi.org/10.1093/bioinformatics/btl446.
44. Gardner SN, Slezak T, Hall BG. 2015. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. Bioinformatics 31:2877–2878. https://doi.org/10.1093/bioinformatics/btv271.
45. Pfeifer B, Wittelsbürger U, Ramos-Onsins SE, Lercher MJ. 2014. PopGenome: an efficient Swiss Army knife for population genomic analyses in R. Mol Biol Evol 31:1929–1936. https://doi.org/10.1093/molbev/msu136.

46. Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586–1591. https://doi.org/10.1093/molbev/msm088.
47. Revel LJ. 2016. Phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol Evol 3:217–223.
48. Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289–290. https://doi.org/10.1093/bioinformatics/btg412.
49. Mostowy R, Croucher NJ, Andam CP, Corander J, Hanage WP, Marttinen P. 2017. Efficient inference of recent and ancestral recombination within bacterial populations. Mol Biol Evol 34:1167–1182. https://doi.org/10.1093/molbev/msx066.
50. Gu Z, Gu L, Eils R, Schlesner M, Brors B. 2014. Circlize implements and enhances circular visualization in R. Bioinformatics 30:2811–2812. https://doi.org/10.1093/bioinformatics/btu393.
51. Gomila M, Moralejo E, Busquets A, Segui G, Olmo D, Nieto A, Jua A, Lalucat J. 2019. Draft genome resources of two strains of *Xylella fastidiosa* XYL1732/17 and XYL2055/17 isolated from Mallorca vineyards. Phytopathol 109:222–224. https://doi.org/10.1094/PHYTO-08-18-0298-A.
52. Denancé N, Legendre B, Briand M, Olivier V, de Boisseson C, Poliakoff F, Jacques M-A. 2017. Several subspecies and sequence types are associated with the emergence of *Xylella fastidiosa* in natural settings in France. Plant Pathology 66:1054–1064. https://doi.org/10.1111/ppa.12695.
53. Bhattacharyya A, Stilwagen S, Reznik G, Feil H, Feil WS, Anderson I, Bernal A, D'Souza M, Ivanova N, Kapatral V, Larsen N, Los T, Lykidis A, Selkov E, Jr, Walunas TL, Purcell A, Edwards RA, Hawkins T, Haselkorn R, Overbeek R, Kyrpides NC, Predki PF. 2002. Draft sequencing and comparative genomics of *Xylella fastidiosa* strains reveal novel biological insights. Genome Res 10:1556–1563. https://doi.org/10.1101/gr.370702.
54. Van Horn C, Chang CJ, Chen J. 2017. *De novo* whole-genome sequence of *Xylella fastidiosa* subsp. *multiplex* strain BB01 isolated from a blueberry in Georgia, USA. Genome Announc 5:e01598-16. https://doi.org/10.1128/genomeA.01598-16.
55. Denance N, Briand M, Gaborieau R, Gaillard S, Jacques MA. 2019. Identification of genetic relationships and subspecies signatures in *Xylella fastidiosa*. BMC Genomics 20:239. https://doi.org/10.1186/s12864-019-5565-9.