e-Xtra*

Multilocus Sequence Typing of *Xylella fastidiosa* Causing Pierce's Disease and Oleander Leaf Scorch in the United States

Xiaoli Yuan, Lisa Morano, Robin Bromley, Senanu Spring-Pearson, Richard Stouthamer, and Leonard Nunney

First, fourth, and sixth authors: Department of Biology, University of California, Riverside 92521; second author: Department of Natural Sciences, University of Houston–Downtown, Houston 77002; and third and fifth authors: Department of Entomology, University of California, Riverside.

Accepted for publication 24 February 2010.

ABSTRACT

Yuan, X., Morano, L., Bromley, R., Spring-Pearson, S., Stouthamer, R., and Nunney, L. 2010. Multilocus sequence typing of *Xylella fastidiosa* causing Pierce's disease and oleander leaf scorch in the United States. Phytopathology 100:601-611.

Using a modified multilocus sequence typing (MLST) scheme for the bacterial plant pathogen *Xylella fastidiosa* based on the same seven housekeeping genes employed in a previously published MLST, we studied the genetic diversity of two subspecies, *X. fastidiosa* subsp. *fastidiosa* and *X. fastidiosa* subsp. *sandyi*, which cause Pierce's disease and oleander leaf scorch, respectively. Typing of 85 U.S. isolates (plus one from northern Mexico) of *X. fastidiosa* subsp. *fastidiosa* from 15 different plant hosts and 21 isolates of *X. fastidiosa* subsp. *sandyi* from 4 different hosts in California and Texas supported their subspecific status. Analysis using the MLST genes plus one cell-surface gene showed no significant genetic differentiation based on geography or host plant within either subspecies. Two cases of homologous recombination (with *X. fastidiosa* subsp. *multiplex*, the third U.S. subspecies) were detected in *X. fastidiosa* subsp. *fastidiosa*. Excluding recombination, MLST site polymorphism in *X. fastidiosa* subsp. *fastidiosa* (0.048%) and *X. fastidiosa* subsp. *sandyi* (0.000%) was substantially lower than in *X. fastidiosa* subsp. *multiplex* (0.240%), consistent with the hypothesis that *X. fastidiosa* subspp. *fastidiosa* and *sandyi* were introduced into the United States (probably just prior to 1880 and 1980, respectively). Using whole-genome analysis, we showed that MLST is more effective at genetic discrimination at the specific and subspecific level than other typing methods applied to *X. fastidiosa*. Moreover, MLST is the only technique effective in detecting recombination.

Understanding the population genetics of bacterial pathogens has become increasingly important because it is a prerequisite for understanding how these pathogens evolve and how genetic changes have allowed them to adapt to different hosts and environments. Molecular characterization of pathogens, coupled with biological tests on host range and ecological parameters, have proven to be very useful in the management of drug resistance, for development of vaccines, and for understanding the epidemiology of the diseases they cause (30,34). Although several genomes are now becoming available for single species, understanding evolutionary relationships requires the characterization of a much larger representative sample of isolates (12).

Classical bacteriological techniques, such as gram staining, sugar utilization analysis, and serological testing, have provided means to identify isolates during short-term outbreaks; however, such methods are inadequate for evolutionary studies (14,36, 46,50). Advances in molecular microbiology have opened new avenues for the phylogenetic analysis of bacterial isolates (27). These include nonsequence-based methods such as restriction fragment length polymorphisms (RFLPs), randomly amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), and simple-sequence repeats (SSRs or microsatellites), plus sequence-based methods targeting specific genes or regions such as 16S rDNA and the 16S-23S internally transcribed spacer (ITS). However, there are several drawbacks to these methods: a RAPD analysis is limited to presence or absence data and reproducibility is poor; the interpretation of genetic information produced by the traditional RFLP analysis and even by the more recent AFLP approach is limited and has largely been replaced by SSRs; SSRs are generally reproducible but the typical high rate of evolution of SSR generally limits their use to problems involving the very recent ancestry of strains; RAPDs, RFLPs, AFLPs, and SSRs all suffer from the problem of the potential homoplasy of alleles that can bias estimated relationships; whereas, conversely, 16S rDNA is too conserved to make inferences about the relatedness of highly similar microorganisms. The 16S-23S ITS generally has a more optimal range of variability but, in common with the other methods, is not generally suitable for detecting recombination between isolates as a source of evolutionary change. More recently, multilocus sequence typing (MLST) has become the method of choice for studying closely related microorganisms because of its high resolution, ease of comparison between laboratories, and suitability for web-based data sharing (26,31,32,35), as well as its ability to recognize recombination between different lines (13,15,58).

MLST characterizes isolates of bacteria on the basis of sequence variation, using approximately 500 bp of nucleotide sequence from internal fragments of seven housekeeping loci (31,32,57). The different sequences at each locus are assigned different allele numbers, and each strain is defined by the alleles at the seven loci (the allelic profile). Each unique allelic profile defines a sequence type (ST), which is a convenient and unambiguous descriptor for strains (1,14), while the complete sequence information is always available for more detailed analysis. This initial assessment of relationships has proven extremely useful in categorizing and tracking the variation and evolution seen in human pathogens (18,33,40,61). MLST employs housekeeping genes to minimize the effects of transient episodes of

Corresponding author: L. Nunney; E-mail address: Leonard.Nunney@ucr.edu

*The *e*-Xtra logo stands for "electronic extra" and indicates that the online version contains a table showing the *X. fastidiosa* isolates used in this study.

doi:10.1094/PHYTO-100-6-0601

selection, a strategy that has been shown to provide high discriminatory power while retaining signatures of longer-term evolutionary relationships and clonal stability (11,33).

Despite broad application of MLST to many human pathogens, the approach has been adopted less rapidly in the study of plant pathogens; however, Sarkar and Guttman (49) developed MLST for *Pseudomonas syringae* and the approach has proven productive in later research (25,64,67). More recently, Danet et al. (9) developed an MLST scheme to characterize fruit tree phytoplasmas.

Scally et al. (50) introduced an MLST system for the plant pathogen *Xylella fastidiosa*. *X. fastidiosa* is a New World γ-proteobacterium that infects the xylem of a wide range of plant hosts (23,24). It typically causes leaf scorch diseases and is spread by xylem-feeding insects, generally leaf hoppers (45). A number of economically important plant species are affected, including grape, oleander, almond, peach, coffee, and citrus (23).

Phylogenetic studies suggested the division of *X. fastidiosa* into four subspecies (51,52): (i) *X. fastidiosa* subsp. *fastidiosa*, the cause of Pierce's disease in grape; (ii) proposed *X. fastidiosa* subsp. *sandyi*, which causes oleander leaf scorch; (iii) *X. fastidiosa* subsp. *multiplex*, associated with scorch disease in a range of trees, including almond, peach, oak, and others; and (iv) proposed *X. fastidiosa* subsp. *pauca*, which causes citrus variegated chlorosis and coffee leaf scorch. Although each subspecies is found in multiple plant hosts (for example, *X. fastidiosa* subsp. *fastidiosa* not only infects grape; it also causes alfalfa dwarf and overlaps with *X. fastidiosa* subsp. *multiplex* in causing almond leaf scorch), there are characteristic subspecific differences in plant hosts (20,51,52), illustrating the importance of always discriminating isolates at the subspecific level. Moreover, there appears to be additional host specificity within some of the subspecies (29,53) that needs further genetic study. Multilocus typing is critically important in documenting host-specific pathogenicity both between and within subspecies, providing the basic genetic information necessary to understand the processes involved in its evolution, and identifying the emergence and spread of novel disease-causing genotypes (50). Scally et al. (50) and Almeida et al. (3) showed that the MLST system for this species was effective in identifying subspecies and, in some cases, plant-host-related subgroups within subspecies, for both North and South American isolates. Furthermore, in both studies, MLST proved effective at recognizing the importance of recombination in the evolution of this species.

MLST in *X. fastidiosa* is complemented by the availability of four fully annotated and two draft genome sequences: the complete sequences of 9a5c (*X. fastidiosa* subsp. *pauca*, isolate CVC0018; Supplementary Table 1) from citrus (56), Temecula-1 (*X. fastidiosa* subsp. *fastidiosa*, isolate PD0001) from grape (63), M12 (*X. fastidiosa* subsp. *multiplex*, isolate ALS0299) and M23 (*X. fastidiosa* subsp. *fastidiosa*, isolate ALS0300) from almond (7), and the draft sequences of Dixon (*X. fastidiosa* subsp. *multiplex*, isolate ALS0003) from almond and Ann1 (*X. fastidiosa* subsp. *sandyi*, isolate OLS0002) from oleander (5). These multiple genomes, combined with an effective and reliable strain typing system such as MLST, provides a solid basis for the development of epidemiological and phylogenetic studies within the *X. fastidiosa* group (10,55).

The first goal of the present study was to refine details of the MLST system to increase its reliability and ensure its applicability to non-U.S. forms (specifically, *X. fastidiosa* subsp. *pauca*) (3). In addition, we used genome sequence data to compare the effectiveness of MLST relative to two other genetic typing methods that have been used to study *X. fastidiosa*: variation in SSRs (28) and in rDNA (19). The second goal was to investigate the genetic variability of two of the taxa, *X. fastidiosa* subspp. *fastidiosa* and *sandyi*, across the United States and across different plant hosts in order to test the hypothesis that the lack of genetic variation previously observed in a small sample of *X. fastidiosa* subspp. *fastidiosa* and *sandyi* isolates (52) reflected strong selection driven by the severe constraints of host plant adaptation. Schuenzel et al. (52) estimated that these taxa have been evolving independently for more than 15,000 years, and yet little variation was found within either subspecies. The MLST analysis of the same data by Scally et al. (50) quantified this homogeneity by showing that each subspecies was defined by a single genotypic grouping (referred to as a clonal complex). In contrast, the other U.S. subspecies, *X. fastidiosa* subsp. *multiplex*, consisted of three distinct clonal complexes. To test this hypothesis, we examined, where possible, isolates collected from a larger number of host plants and a wider geographical range than the previous study to determine if this revealed additional genetic variability.

The third goal was to investigate recombination in *X. fastidiosa* subspp. *fastidiosa* and *sandyi*. Scally et al. (50) and Almeida et al. (3) noted the potentially important role of recombination in generating variation in *X. fastidiosa* but we were particularly interested in the role of intersubspecific recombination, because this may provide an important source of novel genetic variation. Schuenzel et al. (52) noted one case of apparent intersubspecific recombination into an *X. fastidiosa* subsp. *fastidiosa* isolate (PD0014) but did not examine it in detail. We looked for additional examples in our larger set of isolates.

MATERIALS AND METHODS

Isolates. The DNA from 86 isolates of *X. fastidiosa* subsp. *fastidiosa* and 21 isolates of *X. fastidiosa* subsp. *sandyi* were used in this study (Supplementary Table 1). Except for 1 isolate from Baja California, Mexico, all of the other 85 isolates were sampled from symptomatic plants across the United States, mostly from California (69%), Texas (14%), and Florida (13%). An additional eight isolates were included (Supplementary Table 1) as representative examples of *X. fastidiosa* subsp. *multiplex* (six isolates) and *X. fastidiosa* subsp. *pauca* (two isolates). Most of these outgroup isolates have been previously analyzed (50,52), with the exception of the newly sequenced *X. fastidiosa* subsp. *multiplex* genome M12 isolate ALS0299 (7), and PLP0070 and CVC0239, which were added to provide additional diversity. Inclusion of a more extensive MLST data set from *X. fastidiosa* subspp. *multiplex* and *pauca* does not alter any of the patterns discussed (data not shown).

MLST system modification. We modified the *X. fastidiosa* MLST scheme of Scally et al. (50) but retained the same seven housekeeping genes (Table 1). Sequence data from the six available genomes were used to find conserved primer sites that would allow reliable amplification of each target region from isolates collected from all geographic areas. We also ensured that the size of the amplified sequence was small enough for accurate sequencing in a single read. Oligo 6 (48) was used to help design specific primers. The ratio of nonsynonymous to synonymous substitutions in the redesigned MLST amplicons was estimated and the possibility of positive selection acting at some codons was tested using PAML version 4.3 (68), comparing models 1 and 2. Due to the redesign of the primers, some of the previously recognized variation was no longer included in the MLST system; therefore, some alleles from Scally et al. (50) no longer exist. Those vacated allele numbers were reassigned.

Scally et al. (50) sequenced part of the surface protein gene, *pilU*, and found that it revealed novel variation not seen in the MLST genes. For this reason, the gene was included in the present study, thus increasing the diversity of the types of gene monitored; however, it is not part of the MLST scheme and was not included in the allelic profiles.

We used BLAST to ensure that the genes used (the seven MLST genes plus *pilU*) were present as only a single copy in all

X. fastidiosa subspecific clades of *X. fastidiosa* subsp. *fastidiosa* (Temecula-1 and M23), *X. fastidiosa* subsp. *multiplex* (M12 and Dixon), *X. fastidiosa* subsp. *sandyi* (Ann-1), and *X. fastidiosa* subsp. *pauca* (9a5c), all of which are available on the GenBank microbial genomes BLAST website.

PCR amplification and sequence analysis. Each mixture contained template DNA at 1 to 10 ng/μl, 1.25 U of Dream*Taq* and 1× Dream*Taq* Buffer (Fermentas, Burlington, Ontario, Canada), 0.2mM dNTPs (Fermentas), 0.3 μM forward and the reverse primers, and water for a final reaction volume of 25 μl. The following thermocycler (MasterCycler ep; Eppendorf, Hamburg, Germany) program was used: an initial denaturation at 95°C for 3 min; followed by 35 cycles of 95°C for 30 s, 65°C for 30 s, and 72°C for 1 min; with a final extension at 72°C for 10 min. Amplification was confirmed on a 1% agarose gel. Polymerase chain reaction (PCR) products were prepared for sequencing with the Wizard PCR Prep DNA purification kit (Promega Corp., Madison, WI). Sequencing was done on an ABI 3730xl DNA Analyzer (Applied Biosystems, Foster City, CA) by the UCR Core Instrument Facility. Ambiguities between the forward and reverse sequences were resolved be resequencing. Sequences were aligned by eye with BioEdit v7.0.9.

Identification of clonal complexes. Under the MLST methodology, each allele of a gene is given a different number (31). Thus, each *X. fastidiosa* isolate was characterized by its allelic profile, consisting of the seven numbers defining the allele at each of the seven loci. Each unique allelic profile was assigned an ST number. The STs were grouped into clonal complexes using the eBURST (v3) program (14) using the group criterion employed by Scally et al. (50) so that, within each complex, the STs must share five or more alleles with at least one other ST of the clonal complex.

Genetic and phylogenetic analysis. The nucleotide polymorphism and nucleotide diversity within *X. fastidiosa* subsp. *fastidiosa* and *X. fastidiosa* subsp. *sandyi* were estimated based on the polymorphic sites found in the seven MLST genes using DNAsp 4.50.3 (47). To investigate genetic differentiation using a priori groups based on geography and host plant, the genetic variation between populations relative to within-population variation (K_{st}) was calculated using DNAsp (47), which calculates significance using Monte Carlo simulations. To increase statistical power for this test, the *pilU* sequences were added to the MLST data set. This same enlarged dataset was used to estimate Tajima's D (60) again using DNAsp, where significance is estimated following Tajima's (60) assumption of a β distribution.

A maximum likelihood (ML) phylogenetic tree was created for each of the eight genes (the seven MLST plus *pilU*) using all alleles observed, as well as for the concatenated data set using all unique genotypes. This was done both with and without the alleles and isolates identified a priori as derived from inter-subspecific recombination. The evolutionary model used was the ML general time-reversible (GTR) model with γ distributed rate variation using PAUP (59), based on our determination that this model never performed significantly worse than the best-fit model defined using FindModel (http://www.hiv.lanl.gov/content/sequence/findmodel/findmodel.html) (data not shown). Bootstrap values were based on 1,000 replicates. To establish the robustness of the concatenated trees, we also used the ML (Dnaml), and parsimony (Pars) models implemented in Phylip (17). In cases where a best-fit gene tree topology differed from the known subspecific tree (52), the congruence of the observed tree to the expected one was tested using the Kishino-Hasegawa-Templeton (KHT) test in Dnaml.

Genetic distances among the Temecula-1, M23, M12, and 9a5c genomes. We measured the genetic distance among the published genomes of two *X. fastidiosa* subsp. *fastidiosa* strains (Temecula-1 and M23), an *X. fastidiosa* subsp. *multiplex* strain (M12), and one *X. fastidiosa* subsp. *pauca* strain (9a5c) using four different types of DNA often used in bacterial phylogenetic studies: 16S rDNA, the 16S-23S ITS rDNA, protein coding genes (the MLST genes), and SSRs using 22 known loci (28). For the first three types of DNA, genetic distance was based on sequence divergence and calculated using the Phylip program Dnadist using F84 distance (17). For the SSR data, because only pairs of haploid genomes were compared, the genetic distance was simply the average over loci of some measure of the difference in allele size per locus (d). We employed two measures, one assuming stepwise mutation that uses the absolute difference (54) and a second assuming a geometric distribution of mutation size, where the difference is measured as $1 - 2^{-|d|}$ (6).

Nucleotide sequence accession numbers. The gene sequences for the MLST alleles analyzed in this study are available at GenBank under the following accession numbers: FJ610156 to FJ610160, FJ610163, FJ610165 to FJ610169, FJ610171 to

TABLE 1. Polymerase chain reaction (PCR) primers used in the *Xylella fastidiosa* multilocus sequence typing (MLST), including the genomic annotation, function, and amplicon lengths for the seven MLST housekeeping genes used, plus similar detail for the non-MLST cell-surface gene that was also monitored

	Genome annotation[a]				Function			
Gene	9a5c	M12	Tem-1	M23	Gene	Biochemical	Primer sequences[b]	L (bp)[c]
MLST								
holC	XF0136	Xfasm12_0112	PD0104	XfasM23_0096	DNA polymerase III holoenzyme, chi subunit	Replication	for, 5′ ATGGC ACGCG CCGAC TTCT 3′ rev, 5′ ATGTC GTGTT TGTTC ATGTG CAGG 3′	379
nuoL	XF0316	Xfasm12_0280	PD0259	XfasM23_0251	NADH-ubiquinone oxidoreductase, NQO12 subunit	Aerobic respiration	for, 5′ TAGCG ACTTA CGGTT ACTGG GC 3′ rev, 5′ ACCAC CGATC CACAA CGCAT 3′	557
gltT	XF0656	Xfasm12_1656	PD1516	XfasM23_1603	Glutamate symport protein	Transport of amino acids	for, 5′ TCATG ATCCA AATCA CTCGC TT 3′ rev, 5′ ACTGG ACGCT GCCTC GTAAA CC 3′	654
cysG	XF0832	Xfasm12_2018	PD1840	XfasM23_1941	Siroheme synthase	Biosynthesis of heme, porphyrin	for, 5′ GCCGA AGCAG TGCTG GAAG 3′ rev, 5′ GCCAT TTTCG ATCAG TGCAA AAG 3′	600
petC	XF0910	Xfasm_1943	PD1775	XfasM23_1878	Ubiquinol cytochrome C oxidoreductase, C1 subunit	Electron transport	for, 5′ GCTGC CATTC GTTGA AGTAC CT 3′ rev, 5′ GCACG TCCTC CCAAT AAGCC T 3′	533
leuA	XF1818	Xfasm_1205	PD1047	XfasM23_1111	2-Isopropylmalate synthase	Amino acid biosynthesis	for, 5′ GGTGC ACGCC AAATC GAATG 3′ rev, 5′ GTATC GTTGT GGCGT ACACT G 3′	708
malF	XF2447	Xfasm12_1606	PD1465	XfasM23_1549	ABC transporter sugar permease	Transport of carbohydrates	for, 5′ TTGCT GGTCC TGCGG TGTTG 3′ rev, 5′ GACAG CAGAA GCACG TCCCA GAT 3′	730
Non-MLST								
pilU	XF1632	Xfasm12_1254	PD0148	XfasM23_1222	Twitching motility protein	Surface structures	for, 5′ CCGTA ATCAC AACTC AACAG GACA 3′ rev, 5′ CTGCG AATCA GCATG GCGTA 3′	545

[a] Genomes: 9a5c (*X. fastidiosa* subsp. *pauca*), M12 (*X. fastidiosa* subsp. *multiplex*), Temecula-1 (Tem-1) (*X. fastidiosa* subsp. *fastidiosa*), and M23 (*X. fastidiosa* subsp. *fastidiosa*).
[b] Primers used for PCR and for sequencing.
[c] Amplicon length (L) with primers removed.

FJ610176, FJ610179, FJ610181 to FJ610185, FJ610189, FJ610193 to FJ610198, FJ610201, FJ610204 to FJ610207, FJ610210, FJ610213 to FJ610218, FJ610221, and FJ965544 to FJ965546.

RESULTS

Modified MLST. The MLST scheme used is a modified version of the one developed for *X. fastidiosa* by Scally et al. (50), using the same seven genes: *leuA, petC, malF* (originally named *lacF*), *cysG, holC, nuoL*, and *gltT*. However, the specific regions of these genes used was modified either because the original primers gave poor to no amplification with *X. fastidiosa* subsp. *pauca* (3), the South American subspecies, or because excessive length of the original amplicon resulted in poor sequence reads. All of the new primers were effective with *X. fastidiosa* subsp. *pauca* and the maximum amplicon length in the modified MLST was reduced to 730 bp (Table 1). The modifications reduced the combined sequence length of all MLST genes from 6,675 to 4,161 bp.

MLST is based on monitoring housekeeping genes in order to minimize large localized genetic changes due to strong natural selection. We confirmed that, as reported in Scally et al. (50), there was no indication that any of the gene regions targeted were subject to positive selection because the overall ratio of non-synonymous to synonymous substitution (dn/ds) was 0.22, well below the positive selection threshold of 1, and no codons exhibited evidence of positive selection (χ_1^2 = 0.95 ns for accepting that all dn/ds ≤ 1).

We resequenced all of the isolates included in the study of Scally et al. (50); however, the reduction in length and shifted primer positions resulted in the disappearance of some alleles previously identified, because the region in which they were unique was no longer being sequenced. Those allele numbers were reassigned. The alleles from Table 3 of Scally et al. (50) eliminated were as follows: *leuA* allele 4 became equivalent to 1 (ALS0011), *nuoL* allele 4 became 1 (PD0016), allele 5 became 2 (OLS0002), allele 6 became 2 (OLS0009), allele 7 became 3 (ALS0003, OAK0017, OAK0023, OAK0024, PP0027, and PP0028), and *gltT* allele 6 became 5 (PP0028).

We BLASTed the sequences of the seven MLST genes and the one additional non-MLST gene sequenced (*pilU*) with the available genome sequences to confirm that the MLST gene fragments were not duplicated anywhere in the genome of any of the *X. fastidiosa* subspp. They were confirmed to be present as a single copy in Temecula-1 (*X. fastidiosa* subsp. *fastidiosa*: PD0001), M23 (*X. fastidiosa* subsp. *fastidiosa*: ALS0300), M12 (*X. fastidiosa* subsp. *multiplex*: ALS0299), Dixon (*X. fastidiosa* subsp. *multiplex*: ALS0003), and 9a5c (*X. fastidiosa* subsp. *pauca*: CVC0018); however, the Ann1 database revealed duplicates for all of the eight genes that showed 96 to 98% sequence similarity. In all cases, one copy was identical to our MLST sequences for the Ann1 isolate (*X. fastidiosa* subsp. *sandyi*: OLS0002) and the other identical to Dixon (*X. fastidiosa* subsp. *multiplex*: ALS0003). In seven cases, the target sequence was present on two different contigs (leuA contigs 220 and 257, petC 256 and 166, malF 147 and 261, cysG 258 and 226, holC 194 and 252, gltT 261 and 236, and pilU 264 and 257, listing Ann1 and Dixon copies, respectively) whereas, in the case of *nuoL*, the two copies were present on contig 266. The most likely explanation is that the shotgun library of Ann1 is contaminated with an *X. fastidiosa* subsp. *multiplex* genome similar to the Dixon isolate.

MLST analysis. The allelic profiles of each ST are shown in Table 2 and the ST of each isolate is shown in Supplementary Table 1. The data are also available at the MLST website (http://pubmlst.org). The MLST of the 86 isolates of *X. fastidiosa* subsp. *fastidiosa* revealed minor genetic variation, defining only four STs grouped as a single clonal complex (CC1). Of these isolates, 86.0% (74 isolates) had the same ST (ST1), which was designated as the ancestral type. ST2 (eight isolates) and ST3 (one isolate) differed by only 1 bp from ST1. The remaining three isolates defined another ST (ST4), which only differed from ST1 at *cysG*; however, the difference was a much larger 11 single-nucleotide proteins (SNPs) (see below for further analysis of this *cysG* allele). The 21 isolates of *X. fastidiosa* subsp. *sandyi* formed a second genetically uniform clonal complex (CC2) made up of a single ST, ST5. In the case of *X. fastidiosa* subspp. *multiplex* and. *pauca*, no comprehensive analysis was attempted but the examples that were included represented the range of variability of those subspecies (*unpublished data*). They clustered as expected: the *X. fastidiosa* subsp. *multiplex* isolates grouped together into the three clonal complexes (CC3, CC4, and CC5) previously identified (50), and the two isolates of *X. fastidiosa* subsp. *pauca* from Brazil (CVC0018 and COF0239) grouped together as a pair of singleton STs.

Based on the MLST sequence data, the nucleotide polymorphism and nucleotide diversity in *X. fastidiosa* subsp. *fastidiosa*

TABLE 2. Allelic profiles for sequence types (STs) found in the multilocus sequence typing of 86 *Xylella fastidiosa* subsp. *fastidiosa* isolates and 21 *X. fastidiosa* subsp. *sandyi* isolates, plus representative samples of *X. fastidiosa* subspp. *multiplex* and *pauca*[a]

ST[b]	*leuA*	*petC*	*malF*	*cysG*	*holC*	*nuoL*	*gltT*	Occurrence
CC1								
1	1	1	1	1	1	1	1	74
2	1	1	4	1	1	1	1	8
3	1	1	1	20	1	1	1	1
4	1	1	1	4	1	1	1	3
CC2								
5	2	2	2	2	2	2	2	21
CC3								
6	3	3	3	3	3	3	3	1
7	3	3	3	7	3	3	3	1
CC4								
9	3	3	5	5	4	3	4	2
CC5								
10	5	4	3	3	6	3	5	1
26	5	3	3	3	6	3	5	1
Singleton								
13	7	6	7	9	10	7	8	1
14	8	8	8	11	12	9	9	1

[a] Numbering of the alleles within each locus is arbitrary. It is not determined by genetic similarity.
[b] Clonal complex (CC)1 = *X. fastidiosa* subsp. *fastidiosa*; CC2 = *X. fastidiosa* subsp. *sandyi*; CC3, CC4, and CC5 = *X. fastidiosa* subsp. *multiplex*; and singleton type = *X. fastidiosa* subsp. *pauca*.

was 0.048 and 0.005%, respectively (excluding the two highly divergent recombinant alleles discussed below), and, in *X. fastidiosa* subsp. *sandyi*, there was no polymorphism (i.e., both measures were 0.000%). Contrasted to the level of nucleotide polymorphism seen in the representative sample of *X. fastidiosa* subsp. *multiplex* of 0.240%, *X. fastidiosa* subsp. *fastidiosa* and *X. fastidiosa* subsp. *sandyi* exhibited little variation.

Genetic differentiation. Given the low level of genetic variation observed in *X. fastidiosa* subspp. *fastidiosa* and *sandyi*, we added sequence data from the cell-surface gene *pilU* to increase the power of our search for genetic differentiation within the subspecies. In *X. fastidiosa* subsp. *fastidiosa*, although there was one common allele (82%), eight other alleles were found, all at a frequency of less than 5%; in *X. fastidiosa* subsp. *sandyi*, there was a single isolate (OLS0009) with a novel allele which was the only genetic variant found in this group across all eight loci (totaling 4,706 bp).

Comparisons among the three North American subspecies confirmed their highly significant genetic differentiation (all K_{st} > 0.8) (Table 3). The ML phylogenetic tree of the concatenated sequences confirmed this clear differentiation (with ≥ 99% bootstrap support) and reaffirmed the relationships among the subspecific groupings (52) (Fig. 1A). This tree was robust to methodology: the same tree topology was obtained using two different ML methods (PAUP GTR + γ model and Phylip) and using parsimony.

The same subspecific relationships were supported by the phylogenies derived from six of the eight individual genes;

TABLE 3. Genetic differentiation within and between strains using the multilocus sequence typing plus pilU sequence data

Test type[a]	Comparison[b]	K_{st}[c]
Geographical variation within groups		
Xylella fastidiosa subsp. *fastidiosa*	West (38) vs. East (14)	0.001
	West (38) vs. South (10)	0.005
	South (10) vs. East (14)	0.000
X. fastidiosa subsp. *sandyi*	West (14) vs. South (7)	0.000
Plant host variation within groups		
X. fastidiosa subsp. *fastidiosa*	Grape (26) vs. Almond (12)	0.013
	Grape (26) vs. Others (24)[d]	0.005
	Almond (12) vs. Others (24)[d]	0.017
X. fastidiosa subsp. *sandyi*	Oleander (18) vs. Others (3)[e]	0.000
Variation between major groups		
	X. fastidiosa subsp. *fastidiosa* vs. *X. fastidiosa* subsp. *sandyi*	0.922***
	X. fastidiosa subsp. *fastidiosa* vs. *X. fastidiosa* subsp. *multiplex*	0.831***
	X. fastidiosa subsp. *sandyi* vs. *X. fastidiosa* subsp. *multiplex*	0.962***

[a] Geographical variations: West: California and Baja California, Mexico; East: Florida, Georgia, Kentucky, and North Carolina; South: Texas.
[b] Numbers in parentheses indicate the number of isolates used.
[c] K_{st} = genetic variation between populations relative to within-population variation; *, **, and *** indicate $P < 0.05$, 0.01, and 0.001, respectively. All other comparisons were nonsignificant.
[d] Others: cherry, maple, alfalfa, wild grape (two species), New Zealand Christmas tree, Western redbud, Spanish broom, elderberry, citrus, sweet scent, and lupine.
[e] Others: day lily, magnolia, and jacaranda.

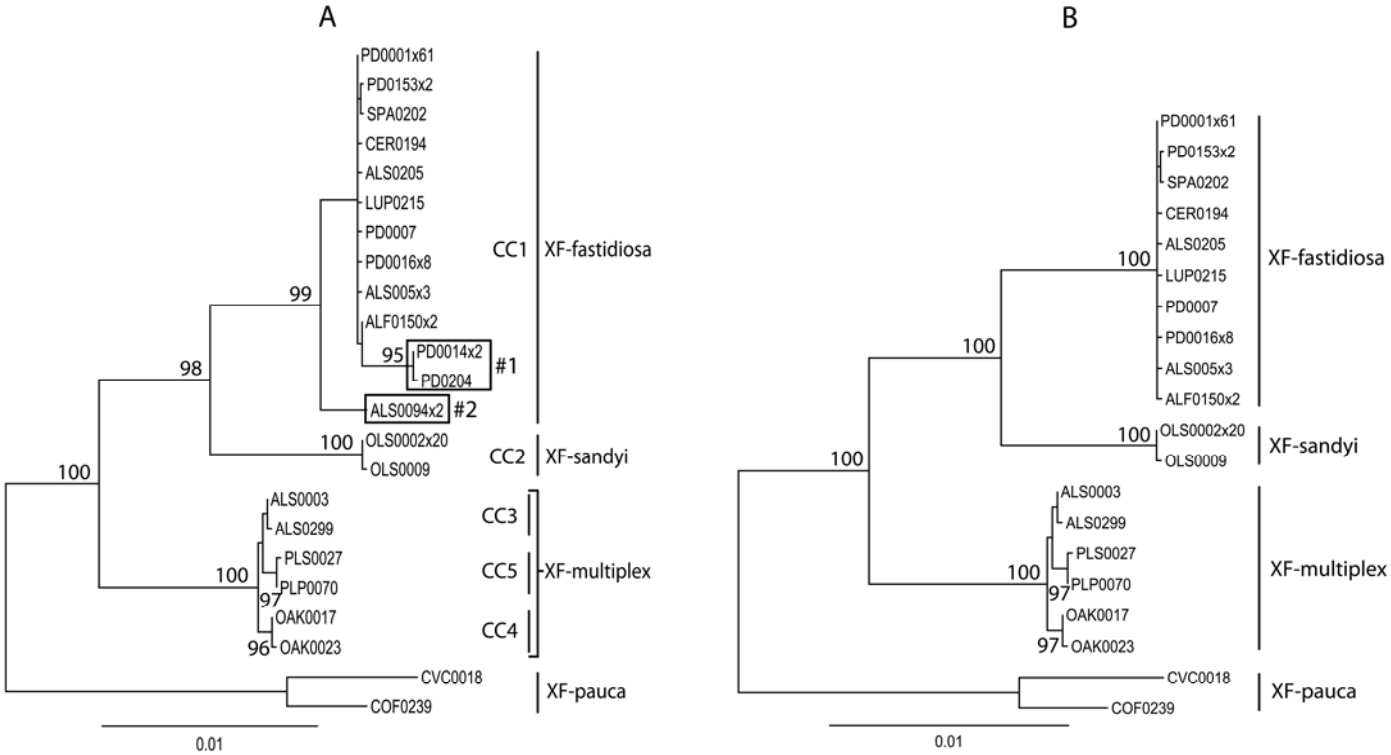

Fig. 1. Maximum likelihood phylogenetic tree of *Xylella fastidiosa* based on the seven multilocus sequence typing loci plus *pilU*. **A,** All genotypes, including those carrying the recombinant *cysG* allele (box no. 1) and *pilU* allele (no. 2). **B,** Nonrecombinant types only. The scale bar defines a 1% sequence divergence. All bootstrap values >80% are shown. Number of isolates (N) with same genotype (i.e., the same sequences across the eight genes) is shown as xN.

however, the trees of two genes showed a single *X. fastidiosa* subsp. *fastidiosa* allele (allele 4 in *cysG* and allele 14 in *pilU*) grouping in the "wrong" position with *X. fastidiosa* subsp. *multiplex* (Fig. 2). This suggested that these two alleles were the product of intersubspecific recombination. These recombinant alleles caused the *cysG* and *pilU* trees to be significantly different from the expected subspecific groupings (see next section). Removing from Figure 1A isolates carrying these recombinant alleles (PD0204/PD0014x2 [i.e., PD0014 plus the identical PD0368, carrying *cysG* allele 4] and ALS0094x2 [i.e., ALS0094 plus the identical ALS0203, carrying pilU allele 14) makes the subspecific pattern more pronounced, and no further subdivision within the *X. fastidiosa* subsp. *fastidiosa* clade or within the *X. fastidiosa* subsp. *sandyi* clade was observed (Fig. 1B).

We investigated two hypotheses that could account for the lack of variation seen in *X. fastidiosa* subsp. *fastidiosa* due to sampling bias. First, we looked for any evidence of geographical substructure. Such substructure, combined with the Californian bias of the available samples, could indicate that variation was being overlooked. However, using all of the available sequence data (MLST plus *pilU*), pairwise comparisons between the west (California), the south (Texas), and the east (Florida, Georgia, Kentucky, and North Carolina) showed no evidence of any geographical genetic structuring that might indicate unsampled variation (Table 3). Second, Schuenzel et al. (52) had suggested that the lack of variation they observed in their limited sample of *X. fastidiosa* subsp. *fastidiosa* isolates (seven from grape and two from almond) could be due to host selection. This hypothesis suggests a reservoir of genetic variation present in other hosts. Analysis comparing isolates from grape (*Vitis vinifera*), almond (*Prunus dulcis*), and other hosts (cherry, *P. avium*; maple, *Acer* sp.; alfalfa, *Medicago sativa*; muscadine, *V. rotundifolia*; desert wild grape, *V. girdiana*; wild grape, *V. aestivalis*; New Zealand Christmas tree, *Metrosideros* sp.; Western redbud, *Cercis occidentalis*; Spanish broom, *Spartium junceum*; sweet scent, *Pluchea odorata*; elderberry, *Sambucus canadensis*; citrus, *Citrus sinensis*; and lupine, *Lupinus aridorum*) showed no differences among them (Table 3), and there was no increase in genetic diversity among isolates from these "other" hosts.

A third hypothesis of a founding bottleneck predicts, in addition to a lack of genetic variation, that variation should show an excess of rare alleles. This bias can be detected as a negative Tajima's D, which measures the difference between nucleotide diversity and a weighted estimate of nucleotide polymorphism. Using the MLST + *pilU* data, Tajima's D was –2.45 ($P < 0.01$).

In the case of *X. fastidiosa* subsp. *sandyi*, of the 21 isolates sequenced, 20 were identical and the 21st differed by only a single base pair. The one variant was found in California on oleander but, given the almost total lack of variation, there was clearly no indication of geographical variation between the Californian and Texan isolates of these subspecies. Similarly, although only three isolates have been obtained from hosts other than oleander (daylily, magnolia, and jacaranda) (Supplementary Table 1), these were genetically identical to all but one of the oleander isolates; therefore, no host-plant differentiation was apparent.

Recombination. Schuenzel et al. (52) suggested that the *cysG* allele of PD0014, isolated in Mendocino County in Northern California, was the result of intersubspecific recombination. We resequenced PD0014 using the modified MLST scheme, confirming the *cysG* sequence as allele 4. We found that two other isolates, PD0204 and PD0368, from nearby Sonoma County carried the same allele. These three isolates defined ST4 (although PD0204 was genetically distinct because it differed at *pilU*) (Supplementary Table 1) and, except at *cysG*, they all carried alleles identical to the abundant *X. fastidiosa* subsp. *fastidiosa* ST1 and grouped within CC1 (Fig. 1). The phylogenetic tree based only on *cysG* (Fig. 2) showed that allele 1 (FAS-1), found in 82 of the 86 *X. fastidiosa* subsp. *fastidiosa* isolates, plus FAS-20 (found in only 1 isolate), together with allele 2 (SAN-2), found in all *X. fastidiosa* subsp. *sandyi* isolates, grouped as expected (97% bootstrap support); however, allele 4, restricted to *X. fastidiosa* subsp. *fastidiosa*, grouped with the *X. fastidiosa* subsp. *multiplex* alleles 3 and 5. This gene tree was

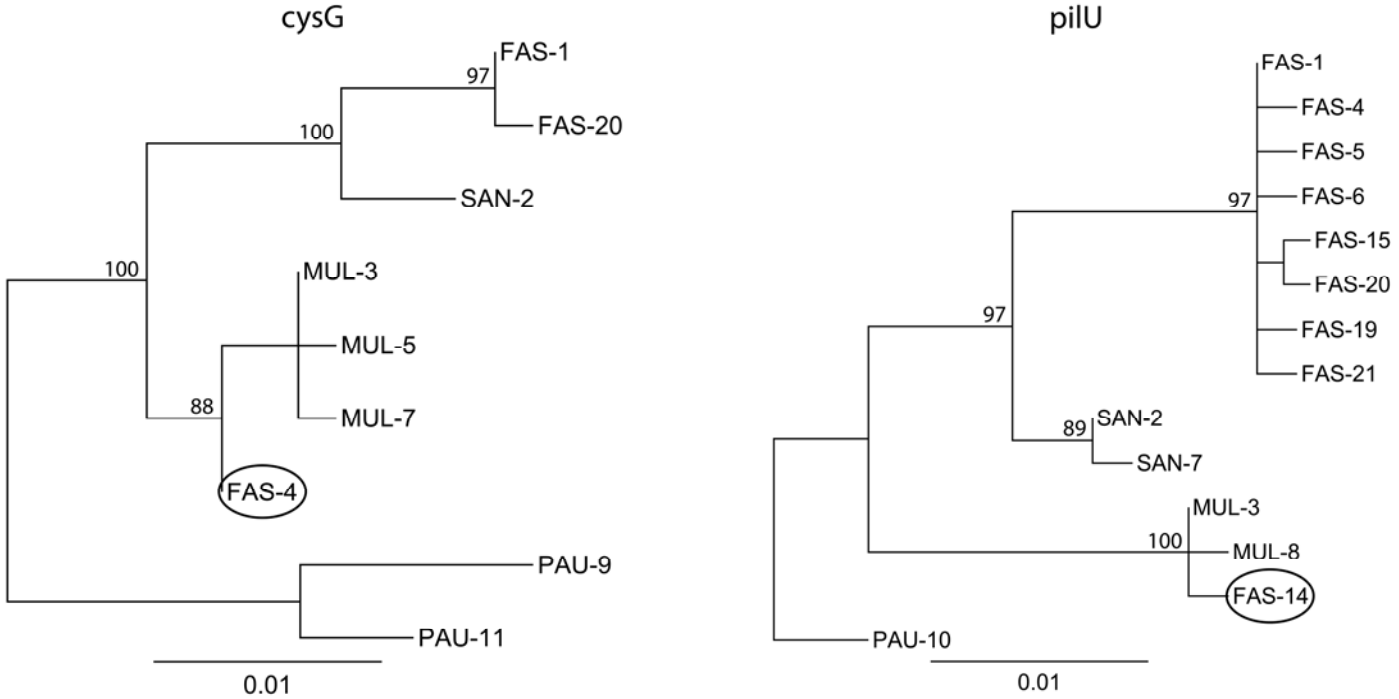

Fig. 2. Single-locus phylogenetic trees for *cysG* and *pilU*. Alleles are labeled according to the subspecies of the isolates in which they were found (FAS, *Xylella fastidiosa* subsp. *fastidiosa*; SAN, XF-sandyi; MUL, *X. fastidiosa* subsp. *multiplex*; PAU, *X. fastidiosa* subsp. *pauca*). Circled alleles indicate alleles of *X. fastidiosa* subsp. *fastidiosa* strains that grouped with *X. fastidiosa* subsp. *multiplex*. All bootstrap values >80% are shown. Scale bar defines a 1% sequence divergence.

significantly different from the known subspecific tree based on the KHT test (Δlog likelihood = 31.5 ± 12.5). Allele 4 differed from the typical *X. fastidiosa* subsp. *fastidiosa* allele 1 at 11 different sites (and from the *X. fastidiosa* subsp. *sandyi* allele 2 at 7 of these sites plus 4 more) but showed identity at all 11 when compared with the *X. fastidiosa* subsp. *multiplex* allele 3, from which it differed at only 2 other sites (Table 4). In the absence of recombination, this pattern requires seven convergent mutations, two nonsynonymous and five synonymous. The involvement of synonymous changes eliminates convergence due to selection and, given seven mutations, the chance of them occurring at seven of the nine sites (out of 545 bp) that distinguish *X. fastidiosa* subsp. *fastidiosa* and *X. fastidiosa* subsp. *sandyi* from *X. fastidiosa* subsp. *multiplex* is $\approx 10^{-14}$ (= 9!/(2!/545^7), without even accounting for the need of each mutation to result in a specific nucleotide. On the other hand, such a pattern is entirely consistent with genetic exchange between *X. fastidiosa* subsp. *multiplex* and *X. fastidiosa* subsp. *fastidiosa*.

A second example of recombination from *X. fastidiosa* subsp. *multiplex* into *X. fastidiosa* subsp. *fastidiosa* was found in the two ST1 isolates, ALS0094 and ALS0203. Because they were ST1, they had the same sequences as the vast majority of *X. fastidiosa* subsp. *fastidiosa* isolates across the MLST genes; however, they each carried a highly divergent allele 14 at *pilU*. This allele grouped with *X. fastidiosa* subsp. *multiplex* with 100% bootstrap support (*pilU*) (Fig. 2), resulting in a tree that was significantly different from the known subspecific tree based on the KHT test (Δlog likelihood = 55.8 ± 15.9). There are 18 nucleotide differences between *pilU* allele 14 and allele 1, and *X. fastidiosa* subsp. *sandyi* allele 2 differs from allele 14 at 12 of these sites plus 2 others (Table 4). In stark contrast, the *X. fastidiosa* subsp. *multiplex* allele 3 differs from allele 14 at only one site, suggesting a recent common ancestry consistent with recombination of sequence from *X. fastidiosa* subsp. *multiplex* into *X. fastidiosa* subsp. *fastidiosa*. More specifically, in the absence of recombination, this pattern requires 11 convergent mutations, 4 nonsynonymous and 7 synonymous. As with cysG, the involvement of synonymous changes eliminates convergence due to selection and, given 11 mutations, the chance of them occurring at the only 11 sites (out of 600 bp) that distinguish *X. fastidiosa* subspp. *fastidiosa* and *sandyi* from *X. fastidiosa* subsp. *multiplex* is $\approx 10^{-23}$ (= 11!/600^{11}), again not accounting for a specific base needing to result from each mutation.

Comparison of genetic typing methods. We compared four different typing methods—multilocus SSR, MLST, 16S-23S ITS sequencing, and 16S sequencing—using the published genomes of *X. fastidiosa* subsp. *fastidiosa* (Temecula-1 and M23), *X. fastidiosa* subsp. *multiplex* (M12), and *X. fastidiosa* subsp. *pauca* (9a5c). The SSR analysis showed that the two *X. fastidiosa* subsp. *fastidiosa* genomes had no alleles in common with the single *X. fastidiosa* subsp. *multiplex* genome, whereas Temecula-1 and M23 differed at 68% (15 of 22) of the SSR loci (Table 5). Based

TABLE 4. Homologous recombination of *Xylella fastidiosa* subsp. *multiplex* sequence into the *X. fastidiosa* subsp. *fastidiosa cysG* and *pilU* genes can be seen by noting the variable sites (single nucleotide polymorphism) present in the *X. fastidiosa* subsp. *multiplex* allele (underlined) that are identical in the recombinant *X. fastidiosa* subsp. *fastidiosa* allele (bold) but differ from the most abundant *X. fastidiosa* subsp. *fastidiosa* (top) and *X. fastidiosa* subsp. *sandyi* (bottom) alleles

X. fastidiosa subsp.	Allele no.	Variable sites in DNA sequence																				
cysG (600 bp)		12	22	47	71	259	266	407	434	435	482	489	496	515	525	560	588					
fastidiosa	1	A	C	G	C	G	T	G	C	G	T	A	C	T	G	A	A					
fastidiosa	**4**	**G**	**C**	**A**	**C**	**A**	**C**	**C**	**T**	**G**	**C**	**G**	**C**	**C**	**G**	**C**	**G**					
multiplex	3	G	T	A	C	A	C	C	T	G	C	G	C	C	A	C	G					
sandyi	2	G	C	G	A	G	T	C	C	A	T	A	T	T	G	C	G					
pilU (545 bp)		48	82	151	187	199	206	285	301	311	313	353	386	402	430	442	465	469	499	500	502	
fastidiosa	1	A	T	A	G	C	C	G	G	C	G	G	G	G	C	G	A	T	T	C	A	
fastidiosa	**14**	**G**	**C**	**G**	**A**	**T**	**T**	**G**	**A**	**A**	**A**	**C**	**G**	**C**	**A**	**A**	**C**	**C**	**C**	**T**	**G**	
multiplex	3	G	C	G	A	T	T	G	A	C	A	C	G	C	A	A	C	C	C	T	G	
sandyi	2	A	T	G	G	C	C	A	G	C	G	G	A	G	C	G	C	C	C	T	G	

TABLE 5. Simple-sequence repeat (SSR) differences among Temecula-1 (*Xylella fastidiosa* subsp. *fastidiosa*: PD0001), M23 (*X. fastidiosa* subsp. *fastidiosa*: ALS0300), and M12 (*X. fastidiosa* subsp. *multiplex*: ALS0209), showing the type and number of repeats found in each genome at each locus

SSR locus[a]	Temecula-1	M23	M12
OSSR-2	$(ATG)_4$	$(ATG)_4$	$(ATG)_3$
OSSR-9	$(TTTCCGT)_{17}$	$(TTTCCGT)_8$	$(TTTCCGT)_3$
OSSR-14	$(TGATCCATCCCTGTG)_6$	$(TGATCCATCCCTGTG)_7$	$(TGATCCATCCCTGTG)_2$
OSSR-16	$(CTGCTA)_8$	$(CTGCTA)_{13}$	$(CTGCTA)_7$
OSSR-17	$(TGCCTG)_8$	$(TGCCTG)_9$	$(TGCCTG)_7$
OSSR-19	$(CAGGATCA)_6$	$(CAGGATCA)_6$	$(CAGGATCA)_7$
OSSR-20	$(AGGATGCTA)_{19}$	$(AGGATGCTA)_{14}$	$(AGGATGCTA)_2$
ASSR-9	$(CAATGAC)_8$	$(CAATGAC)_8$	$(CAATGAC)_4$
ASSR-11	$(ACGCATC)_{10}$	$(ACGCATC)_9$	$(ACGCATC)_{14}$
ASSR-12	$(GATTCAG)_8$	$(GATTCAG)_8$	$(GATTCAG)_9$
ASSR-14	$(CTGCGTGC)_9$	$(CTGCGTGC)_9$	$(CTGCGTGC)_7$
ASSR-16	$(GCTCCGGTTCTA)_{19}$	$(GCTCCGGTTCTA)_{17}$	$(GCTCCGGTTCTA)_{12}$
ASSR-19	$(ACAACG)_4$	$(ACAACG)_4$	$(ACAACG)_6$
ASSR-20	$(ACAGAAA)_{18}$	$(ACAGAAA)_9$	$(ACAGAAA)_5$
GSSR-4	$(ATCC)_{18}$	$(ATCC)_{17}$	$(ATCC)_{12}$
GSSR-6	$(CTTGT)_{14}$	$(CTTGT)_{12}$	$(CTTGT)_6$
GSSR-7	$(GGCAAC)_{24}$	$(GGCAAC)_{25}$	$(GGCAAC)_{11}$
GSSR-12	$(TATCTGT)_{21}$	$(TATCTGT)_{14}$	$(TATCTGT)_2$
GSSR-14	$(TCCCGTA)_{24}$	$(TCCCGTA)_6$	$(TCCCGTA)_5$
GSSR-15	$(AGCCTGC)_{17}$	$(AGCCTGC)_{16}$	$(AGCCTGC)_4$
GSSR-19	$(GAAAACAAG)_{20}$	$(GAAAACAAG)_{21}$	$(GAAAACAAG)_1$
GSSR-20	$(GAACCACTA)_7$	$(GAACCACTA)_7$	$(GAACCACTA)_2$

[a] SSR locus names from Lin et al. (28).

on these data, the genetic distance between the two *X. fastidiosa* subsp. *fastidiosa* genomes was 42% of the distance between the two subspecies using the single-step mutation model and 58% using the geometric model (Table 6). As expected, this closer relationship within (versus between) subspecies was highly significant ($P < 0.001$), based on a sign test: the smallest allelic length difference was the within-subspecies comparison in 20 or 19 of 22 loci (Temecula-1 or M23 versus M12) (Table 5). None of these 22 SSR loci could be located in the *X. fastidiosa* subsp. *pauca* genome; therefore, more distant subspecific comparisons were hypothetical, although the geometric distance measure has a maximum value of 1.0 (Table 6).

MLST analysis showed that Temecula-1 and M23 were both ST1 (although they differed at *pilU* by a single base pair). Whereas the two *X. fastidiosa* subsp. *fastidiosa* genomes differed from M12 by 69 bp, both differed similarly from 9a5c (136 and 127 bp, respectively). Thus, within subspecies, the genetic distance was 0% of the between-subspecies value, and the distance between *X. fastidiosa* subsp. *fastidiosa* and *X. fastidiosa* subsp. *multiplex* was 55% of their average distance to *X. fastidiosa* subsp. *pauca* (Table 6). Based on the 16S-23S spacer, the genetic distances for the three North American strains were very similar to the MLST result: no difference between the two *X. fastidiosa* subsp. *fastidiosa* isolates and 7 SNPs in 513 bp distinguishing *X. fastidiosa* subsp. *fastidiosa* and *X. fastidiosa* subsp. *multiplex*. The 16S-23S difference between the two U.S. subspecies (*X. fastidiosa* subsp. *fastidiosa* and *X. fastidiosa* subsp. *multiplex*) and the South American *X. fastidiosa* subsp. *pauca* was 11 and 4 SNPs, respectively. both measures being lower than the MLST distances. The 16S sequences were also identical within *X. fastidiosa* subsp. *fastidiosa*. The distances between the subspecies were as follows: *X. fastidiosa* subsp. *fastidiosa* versus *X. fastidiosa* subsp. *multiplex* was 0.002, with 3 SNPs in 1,537 bp; *X. fastidiosa* subsp. *fastidiosa* versus *X. fastidiosa* subsp. *pauca* was 0.003, with 4 SNPs; and *X. fastidiosa* subsp. *multiplex* versus *X. fastidiosa* subsp. *pauca* was 0.002, with 3 SNPs (Table 6).

DISCUSSION

In this study, we modified the MLST system to make it more effective and robust in strain typing. With the improved system, we examined the genetic structure of two subspecies of *X. fastidiosa*, one causing Pierce's disease (*X. fastidiosa* subsp. *fastidiosa*) and the other causing oleander leaf scorch (*X. fastidiosa* subsp. *sandyi*). Little genetic variation was found within these two subspecies. This was reflected in the eBURST analysis, which classified both the 86 *X. fastidiosa* subsp. *fastidiosa* isolates and the 21 *X. fastidiosa* subsp. *sandyi* isolates as single clonal complexes (CC1 and CC2, respectively) (Fig. 1A). The DNA polymorphism of the MLST loci across all 86 *X. fastidiosa* subsp. *fastidiosa* isolates and 21 *X. fastidiosa* subsp. *sandyi* was 0.048 and 0.000%, respectively (excluding the recombinant *cysG* allele found in two *X. fastidiosa* subsp. *fastidiosa* isolates), which was a fivefold reduction in genetic variability compared with a representative sample of *X. fastidiosa* subsp. *multiplex* (0.240%). Although *X. fastidiosa* subsp. *sandyi*, at 0 SNPs per 4,161bp, is clearly monomorphic, the level observed in *X. fastidiosa* subsp. *fastidiosa*, roughly 1 SNP per 2 kb, is well within the range (1 SNP per 200 bp) defining genetically monomorphic bacteria (2).

These results confirm the conclusions drawn from the previous analysis of Scally et al. (50), who typed nine isolates of *X. fastidiosa* subsp. *fastidiosa*, seven from grape and two from almond, and six isolates of *X. fastidiosa* subsp. *sandyi*, all from oleander. This lack of variability is not due to the very recent origin of these taxa as specialists on novel host plants, associated with Spanish settlers introducing European grape and oleander, because sequence data indicated, conservatively, a divergence some 15,000 to 25,000 years ago (52). Three other hypotheses might account for the lack of genetic variation of these subspecies. First, strong plant host selection could favor specific clones in specific hosts (52). In the previous study, all *X. fastidiosa* subsp. *fastidiosa* isolates were from grape or almond and all *X. fastidiosa* subsp. *sandyi* isolates were from oleander. Second, geographical structuring within the United States might result in an apparent lack of variability if sampling was focused in a single region. In the previous study, all samples were from California except for one isolate of each subspecies. Third, introduction of the subspecies into the United States from elsewhere could result in a founder effect that would reduce genetic variability. A corollary of this hypothesis is that these subspecies must occur elsewhere outside of the United States, where they exhibit higher levels of genetic variation.

X. fastidiosa subsp. *sandyi* was sampled roughly equally from different places in California and Texas. Despite this geographical separation, the isolates showed complete genetic homogeneity, with the exception of one isolate that had a 1-base difference (in the non-MLST gene, *pilU*) out of 4,706 bases. This divergent genotype was isolated from oleander; therefore, although only three of the isolates were from alternate hosts, they were all identical to each other and to 18 of the 19 oleander isolates. Thus, we found no support for the possibility that there was additional variation in *X. fastidiosa* subsp. *sandyi* within the United States, distributed either geographically or among different plant hosts.

The isolates of *X. fastidiosa* subsp. *fastidiosa* were sampled even more widely across the United States (primarily California, Texas, and Florida) but there was no evidence of significant geographical differentiation. Similarly, comparison of isolates from commercial grape (two species), almond, and other plant hosts (cherry, maple, alfalfa, two species of wild grape, New Zealand Christmas tree, Western redbud, Spanish broom, elderberry, citrus, sweet scent, and lupine) failed to reveal any genetic heterogeneity that might indicate some reservoir of genetic variation in alternate plant hosts. Thus, although *X. fastidiosa* subsp. *fastidiosa* is more genetically variable then *X. fastidiosa* subsp. *sandyi*, the variability is minimal at 1 SNP per 2 kb, and there was no indication that additional variability was distributed geographically across the United States or among different plant hosts.

Because there was no evidence for substantial geographic differentiation or plant-host differentiation within *X. fastidiosa*

TABLE 6. Genetic distances among Temecula-1 (*Xylella fastidiosa* subsp. *fastidiosa*: PD0001), M23 *X. fastidiosa* subsp. *fastidiosa*: ALS0300), M12 (*X. fastidiosa* subsp. *multiplex*: ALS0299), and 9a5c (*X. fastidiosa* subsp. *pauca*: CVC0018) genomes using 16S/16S-23S internal transcribed spacer (above the diagonal) and simple-sequence repeats (SSRs)/multilocus sequence typing genes (below the diagonal)[a]

	Isolate			
Isolate	Temecula-1	M23	M12	9a5c
Temecula-1	…	0.000/0.000	0.002/0.012	0.003/0.020
M23	0.495/0.000	…	0.002/0.012	0.003/0.020
M12	0.852/0.018	0.860/0.018	…	0.002/0.008
9a5c	1.0/0.034	1.0/0.034	1.0/0.031	…

[a] See text for a description of the specific measures used; 1.0 indicates no corresponding SSR loci between North American isolates and South American isolate.

subspp. *fastidiosa* and *sandyi* that could lead to a large underestimate of genetic variability (due to a sampling bias), it becomes probable that the lack of variability in both of these subspecies is real and was due to a founder effect. Under this hypothesis, the greater variability seen in *X. fastidiosa* subsp. *fastidiosa* relative to *X. fastidiosa* subsp. *sandyi* could be due to its earlier introduction into the United States, a pattern consistent with the known history of these pathogens. Pierce's disease of grapevine, the disease caused by *X. fastidiosa* subsp. *fastidiosa*, first appeared in the United States in Southern California in the 1880s (41), whereas oleander leaf scorch, the disease caused by *X. fastidiosa* subsp. *sandyi*, was first observed 100 years later, in the 1980s, also in Southern California (43). The distribution of variation in *X. fastidiosa* subsp. *fastidiosa* is biased toward rare variants (defining a significantly negative Tajima's D statistic), as expected during the reestablishment of variation following a severe bottleneck. Moreover, Montero-Astua (37,38) recently showed that isolates from coffee and grape from Costa Rica were genetically similar to *X. fastidiosa* subsp. *fastidiosa*, raising the possibility that *X. fastidiosa* subsp. *fastidiosa* originated in Central America (L. Nunney, *unpublished data*).

Precise estimation of the timing of these introductions is not possible because the build-up of genetic variability following a founder event depends upon the effective size of the bacterial population; however, the finding of only one isolate carrying a single variant SNP (out of the 4,706 bp examined) in *X. fastidiosa* subsp. *sandyi* is consistent with the recent introduction of a single genotype. Similarly, the level of variation seen in *X. fastidiosa* subsp. *fastidiosa*, while markedly greater than that of *X. fastidiosa* subsp. *sandyi*, again indicates the introduction of a single genotype: the negative Tajima's D supports the occurrence of a strong bottleneck followed by a gradual accumulation of rare alleles and, most notably, all of these alleles were only one mutational step away from the allele found in the dominant genotype of ST1 plus *pilU* allele 1. This genotype made up 60% of the samples despite the widespread geographical and plant-host sampling, and we found no evidence of a second potential ancestral type within the United States. However, *X. fastidiosa* subsp. *fastidiosa* has been present in the United States since at least the 1880s (41), and the reestablishment of genetic variation is expected to be considerably more rapid at SSR (microsatellite) loci due to their much higher mutation rates. Data from the two genomes Temecula-1 and M23 illustrate this recovery: the two genomes differed at 16 of 22 SSR loci (Table 5) but they were identical across all 4,161 bp covered by MLST. At a larger scale, the very low genetic variability seen in the complete MLST dataset of 86 isolates contrasts markedly with the extensive SSR variation detected in a study of 23 *X. fastidiosa* subsp. *fastidiosa* isolates from California (28), consistent with the expectation that SSR variability has recovered from the initial genetic bottleneck.

MLST is extremely useful in identifying and grouping bacterial isolates; however, it can also be used to investigate recombination (13,15,58). Recombination in *X. fastidiosa* was first reported by Scally et al. (50) and, here, we provided further evidence of the homologous recombination of *X. fastidiosa* subsp. *multiplex* alleles into *X. fastidiosa* subsp. *fastidiosa*. The recombinant origin of the *cysG* allele 4 in PD0014 was previously noted (50,52), and we found two more examples of isolates carrying it. This allele appears to have originated by a genetic transfer from *X. fastidiosa* subsp. *multiplex* because allele 4 and the *X. fastidiosa* subsp. *multiplex* allele 3 share 11 base differences from the alleles found in all other isolates of *X. fastidiosa* subsp. *fastidiosa*. Similarly, we found 2 *X. fastidiosa* subsp. *fastidiosa* isolates carrying *pilU* allele 14, which differs from *X. fastidiosa* subsp. *multiplex* allele 3 at a single site, but these two alleles share 17 site differences from the alleles found in the other 60 *X. fastidiosa* subsp. *fastidiosa* isolates.

MLST is one of several potential typing systems. Almeida et al. (3), in their analysis of *X. fastidiosa* subsp. *pauca*, compared MLST with SSR analysis. They concluded that the high variability of SSR loci can be a very powerful tool for studies at the local population level, while MLST is more appropriate for larger-scale investigations. For example, in the study of Lin et al. (28), 23 *X. fastidiosa* subsp. *fastidiosa* and 5 *X. fastidiosa* subsp. *sandyi* isolates were all found to be genetically unique based on 34 SSR loci; however, the power to detect the phylogenetic relationships among *X. fastidiosa* subsp. *fastidiosa*, *X. fastidiosa* subsp. *sandyi*, and *X. fastidiosa* subsp. *multiplex* was lost. We quantified this SSR variability directly from genome sequences (Table 6) and showed that the genetic differences within *X. fastidiosa* subsp. *fastidiosa* averaged 50% of those between *X. fastidiosa* subsp. *fastidiosa* and *X. fastidiosa* subsp. *multiplex*, whereas MLST data showed the differences within subspecies to be less than 1% of those between subspecies. This dramatic difference illustrates how quickly SSR variation builds up and saturates (noting that the *X. fastidiosa* subspp. *fastidiosa* and *multiplex* differed at all of the SSR loci, whereas none of these loci could even be found in the more distantly related *X. fastidiosa* subsp. *pauca*). This saturation distorts the time scale of evolutionary change and strongly supports the conclusion of Almeida et al. (3) that SSRs are only appropriate for intrapopulation studies where there is a very short time scale. For example, SSRs are expected to be very useful in tracking the epidemiological spread of a new and virulent genotype from one season to a next. MLST is more appropriate at longer times scales and is the only method appropriate for detecting recombination.

Another commonly used phylogenetic approach is based on rDNA sequences. Based on our genome comparison, both the 16S and the intergenic spacer gave information roughly equivalent to MLST based on a single locus (rather than seven). Therefore, it is much less powerful than MLST and subject to potentially unrepresentative local genomic events. The 16S gene showed little change at this evolutionary scale, with only a 0.2 and 0.3% difference between subspecies. Clearly, this approach lacks power within subspecies.

In summary, we found that, compared with single-locus analysis, such as using 16S rDNA, which provides too little variation, and SSR, which provides too much variation for evolutionary analysis, the MLST method has clear advantages. MLST facilitates rapid recognition of novel isolates; has proven itself to be very useful in phylogenetic research, including the detection of recombination; and provides an accurate, high-throughput platform for large-scale pathogen typing.

In applying MLST to the problem of detecting and understanding within-species genetic patterns in two of the *X. fastidiosa* subspp., our analysis of 86 *X. fastidiosa* subsp. *fastidiosa* and 21 *X. fastidiosa* subsp. *sandyi* isolates suggests that both subspecies show a lack of genetic variation consistent with their introduction into the United States just prior to the first recorded outbreaks of Pierce's disease in the 1880s (for *X. fastidiosa* subsp. *fastidiosa*) and of oleander leaf scorch in the 1980s (for *X. fastidiosa* subsp. *sandyi*). However, the timing of the introduction of *X. fastidiosa* subsp. *fastidiosa* is somewhat uncertain because of anecdotal evidence that Pierce's disease was the cause of the rapid failure of grapevines planted in the Gulf Coast region in the early 1800s (J. Kamas, *personal communication*). If this anecdotal evidence is correct, the California outbreak resulted from an introduction from the Gulf (22) but this does not alter the evidence pointing to a single successful introduction into the United States. The MLST analysis also enabled us to detect two clear examples of homologous recombination of DNA from *X. fastidiosa* subsp. *multiplex* into *X. fastidiosa* subsp. *fastidiosa*. Such observations of recombination are important because they suggest a mechanism by which *X. fastidiosa* may rapidly adapt to new plant hosts.

ACKNOWLEDGMENTS

This study was supported by a United States Department of Agriculture–Cooperative State Research, Education, and Extension Service–NRI grant 2007-55605-17834 to L. Nunney, R. Stouthamer, and L. Morano. We thank M. Black, J. Hartman, R. Hernandez-Martinez, D. Cooksey, D. Hopkins, E. Schuenzel, N. Schaad, D. Vickerman, F. Wong, J. Hartung, S. Purcell, and R. Almeida for providing the strains used in this study; and L. Enders, J. M. Irving, D. Garcia, C. Adams, D. Raterman, J. Wang, and E. Mah for DNA extraction or sequencing genes of many of the isolates used in this study.

LITERATURE CITED

1. Aanensen, D. M., and Sprat, B. G. 2005. The multilocus sequence typing network: mlst.net. Nucleic Acids Res. 33:W728-33.
2. Achtman, M. 2008. Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. Annu. Rev. Microbiol. 62:53-70.
3. Almeida, R. P. P., Nascimento, F. E., Chau, J., Prado, S. S., Tsai, C. W., Lopes, S. A., and Lopes, J. R. S. 2008. Genetic structure and biology of *Xylella fastidiosa* strains causing disease in citrus and coffee in Brazil. Appl. Environ. Microbiol. 74:3690-3701.
4. Almeida, R. P. P., and Purcell, A. H. 2003. Biological traits of *Xylella fastidiosa* strains from grapes and almonds. Appl. Environ. Microbiol. 69:7447-7452.
5. Bhattacharyya, A., Stilwagen, S., Ivanova, N., D'Souza, M., Bernal, A., Lykidis, A., Kapatral, V., Anderson, L., Larsen, N., Los, T., Reznik, G., Selkov, E., Walunas, T. L., Feil, H., Feil, W. S., Purcell, A., Lassez, J. L., Hawkins, T. L., Haselkorn, R., Overbeek, R., Predki, P. F., and Kyrpides, N. C. 2002. Whole-genome comparative analysis of three phytopathogenic *Xylella fastidiosa* strains. Proc. Natl. Acad. Sci. USA 99:12403-12408.
6. Bruvo, R., Michiels, N. K., D'Souza, T. G., and Schulenburg, H. 2004. A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. Mol. Ecol. 13:2101-2106.
7. Chen, J., Han, S., Civerolo, E., Stenger, D., and Van Sluys, M. 2007. Two whole genome sequences of *Xylella fastidiosa* almond leaf scorch strains. (Abstr.) Phytopathology 97:S21.
8. Costa, H. S., Raetz, E., and Pinckard, T. R. 2004. Plant hosts of *Xylella fastidiosa* in and near Southern California vineyards. Plant Dis. 88:1255-1261.
9. Danet, J. L., Bonnet, P., Jarausch, W., Carraro, L., Skoric, D., Labonne, G., and Foissac, X. 2007. Imp and secY, two new markers for MLST (multi locus sequence typing) in the 16SrX phytoplasma taxonomic group. Bull. Insectol. 60:339-340.
10. Doddapaneni, H., Yao, J. Q., Lin, H., Walker, M. A., and Civerolo, E. L. 2006. Analysis of the genome-wide variations among multiple strains of the plant pathogenic bacterium *Xylella fastidiosa.* BMC Genomics 7:225.
11. Enright, M. C., and Spratt, B. G. 1999. Multilocus sequence typing. Trends Microbiol. 7:482-487.
12. Feil, E. J., and Enright, M. C. 2004. Analyses of clonality and the evolution of bacterial pathogens. Curr. Microbiol. 7:308-313.
13. Feil, E. J., Enright, M. C., and Spratt, B. G. 2000. Estimating the relative contributions of mutation and recombination to clonal diversification: A comparison between *Neisseria meningitidis* and *Streptococcus pneumoniae.* Res. Microbiol. 151:465-469.
14. Feil, E. J., Li, B. C., Aanensen, D. M., Hanage, W. P., and Spratt, B. G. 2004. eBURST: Inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. J. Bacteriol. 186:1518-1530.
15. Feil, E. J., Maiden, M. C. J., Achtman, M., and Spratt, B. G. 1999. The relative contributions of recombination and mutation to the divergence of clones of *Neisseria meningitides*. Mol. Biol. Evol. 16:1496-1502.
16. Feil, H., and Purcell, A. H. 2001. Temperature-dependent growth and survival of *Xylella fastidiosa* in vitro and in potted grapevines. Plant. Dis. 85:1230-1234.
17. Felsenstein, J. 1989. PHYLIP—phylogeny inference package (version 3.2). Cladistics 5:164-166.
18. Gonzalez-Escalona, N., Martinez-Urtaza, J., Romero, J., Espejo, R. T., Jaykus, L. A., and DePaola, A. 2008. Determination of molecular phylogenetics of *Vibrio parahaemolyticus* strains by multilocus sequence typing. J. Bacteriol. 190:2831-2840.
19. Hendson, M., Purcell, A. H., Chen, D., Smart, C., Guilhabert, M., and Kirkpatrick, B. 2001. Genetic diversity of Pierce's Disease strains and other pathotypes of *Xylella fastidiosa*. Appl. Environ. Microbiol. 67:895-903.
20. Hernandez-Martinez, R., de la Cerda, K. A., Costa, H. S., Cooksey, D. A., and Wong, F. P. 2007. Phylogenetic relationships of *Xylella fastidiosa* strains isolated from landscape ornamentals in Southern California. Phytopathology 97:857-864.
21. Hernandez-Martinez, R., Pinckard, T. R., Costa, H. S., Cooksey, D. A., and Wong, F. P. 2006. Discovery and characterization of *Xylella fastidiosa* strains in Southern California causing mulberry leaf scorch. Plant Dis. 90:1143-1149.
22. Hewitt, W. B. 1958. The probable home of Pierce's disease virus. Plant Dis. Rep. 42:211-215.
23. Hopkins, D. L. 1989. *Xylella fastidiosa:* Xylem-limited bacterial pathogen of plants. Annu. Rev. Phytopathol. 27:271-290.
24. Hopkins, D. L., and Purcell, A. H. 2002. *Xylella fastidiosa*: Cause of Pierce's disease of grapevine and other emergent diseases. Plant Dis. 86:1056-1066.
25. Hwang, M. S. H., Morgan, R. L., Sarkar, S. F., Wang, P. W., and Guttman, D. S. 2005. Phylogenetic characterization of virulence and resistance phenotypes of *Pseudomonas syringae*. Appl. Environ. Microbiol. 71:5182-5191.
26. Jolley, K. A., Chan, M. S., and Maiden, M. C. J. 2004. mlstdbNet—distributed multi-locus sequence typing (MLST) databases. BMC Bioinf. 5:86.
27. Lévesque, C. A. 2001. Molecular methods for detection of plant pathogens-what is the future? Can. J. Plant Pathol. 24:333-336.
28. Lin, H., Civerolo, E. L., Hu, R., Barros, S., Francis, M., and Walker, M. A. 2005. Multilocus simple sequence repeat markers for differentiating strains and evaluating genetic diversity of *Xylella fastidiosa.* Appl. Environ. Microbiol. 71:4888-4892.
29. Lopes, S. A., Marcussi, S., Torres, S. C. Z., Souza, V., Fagan, C., Franca, S. C., Fernandes, N. G., and Lopes, J. R. S. 2003. Weeds as alternative hosts of the citrus, coffee and plum strains of *Xylella fastidiosa* in Brazil. Plant Dis. 87:544-549.
30. Louws, F. J., Rademaker, J. L. W., and Bruijn, F. J. 1999. The three D's of PCR-based genomic analysis of phytobacteria: Diversity, detection, and disease diagnosis. Annu. Rev. Phytopathol. 37:81-125.
31. Maiden, M. C. J. 2006. Multilocus sequence typing of bacteria. Annu. Rev. Microbiol.. 60:561-88.
32. Maiden, M. C. J., Bygraves, J. A., Feil, E., Morelli, G., Russell, J. E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D. A., Feavers, I. M., Achtman, M., and Spratt, B. G. 1998. Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. Proc. Natl. Acad. Sci. USA 95:3140-3145.
33. Margos, G., Gatewood, A. G., Aanensen, D. M., Hanincova, K., Terekhova, D., Vollmer, S. A., Cornet, M., Piesman, J., Donaghy, M., Bormane, A., Hurn, M. A., Feil, E. J., Fish, D., Casjens, S., Wormser, G. P., Schwartz, I., and Kurtenbach, K. 2008. MLST of housekeeping genes captures geographic population structure and suggests a European origin of *Borrelia burgdorferi*. Proc. Natl. Acad. Sci. USA 105:8730-8735.
34. Martin, R. R. 2000. Impacts of molecular diagnostic technologies on plant disease management. Annu. Rev. Phytopathol. 38:207-239.
35. Mayor, D., Jores, J., Korczak, B. M., and Kuhnert, P. 2008. Multilocus sequence typing (MLST) of *Mycoplasma hyopneumoniae*: A diverse pathogen with limited clonality. Vet. Microbiol. 127:63-72.
36. Mehta, A., and Rosato, Y. B.. 2001. Phylogenetic relationships of *Xylella fastidiosa* strains from different hosts, based on 16S rDNA and 16S-23S intergenic spacer sequences. Int. J. Syst. Evol. Microbiol. 51:311-318.
37. Montero-Astua, M., Chacon-Diaz, C., Aguilar, E., Rodriguez, C. M., Garita, L., Villalobos, W., Moreira, L., Hartung, J. S., and Rivera, C. 2008. Isolation and molecular characterization of *Xylella fastidiosa* from coffee plants in Costa Rica. J. Microbiol. 46:482-490.
38. Montero-Astua, M., Hartung, J. S., Aguilar, E., Chacon, C., Li, W., Albertazzi, F. J., and Rivera, C. 2007. Genetic diversity of *Xylella fastidiosa* strains from Costa Rica, São Paulo, Brazil, and the United States. Phytopathology 97:1338-1346.
39. Morano, L. D., Bextine, B. R., Garcia, D. A., Maddox, S. V., Gunawan, S., Vitovsky, N. J., and Black, M. C. 2008. Initial genetic analysis of *Xylella fastidiosa* in Texas. Curr. Microbiol. 56:346-351.
40. Perez-Losada, M., Browne, E. B., Madsen, A., Wirth, T., Viscidi, R. P., and Crandall, K. A. 2006. Population genetics of microbial pathogens estimated from multilocus sequence typing (MLST) data. Infect. Genet. Evol. 6:97-112.
41. Pierce, N. B. 1892 The California vine disease. A preliminary report of investigations. U. S. Dep. Agric. Div. Veg. Pathol. Bull. 2. Government Printing Office, Washington, DC.
42. Postinikova, E., Baldwin, C., Whitehouse, C. A., Sechler, A., Schaad, N. W., Sampath, R., Harpin, V., Li, F., Melton, R., Blyn, L., Drader, J., Hofstadler, S., and Schneider, W. L. 2008. Identification of bacterial plant pathogens using multilocus polymerase chain reaction/electrospray ionization-mass spectrometry. Phytopathology 98:1156-1164.
43. Purcell, A. H., Saunders, S. R., Hendson, M., Grebus, M. E., and Henry, M. J. 1999. Causal role of *Xylella fastidiosa* in oleander leaf scorch. Phytopathology 89:53-58.

44. Qin, X. T., Miranda, V. S., Machado, M. A., Lemos, E. G. M., and Hartung, H. S. 2001. An evaluation of the genetic diversity of *Xylella fastidiosa* isolated from diseased citrus and coffee in Sao Paulo, Brazil. Phytopathology 91:599-605.
45. Redak, R. A., Purcell, A. H., Lopes, J. R. S., Blua, M. J., Mizell, R. F., and Andersen, P. C. 2004. The biology of xylem fluid-feeding insect vectors of *Xylella fastidiosa* and their relation to disease epidemiology. Annu. Rev. Entomol. 49:243-270.
46. Rogall, T., Wolters, J., Flohr, T., and Bottger, E. C. 1990. Towards a phylogeny and definition of species at the molecular level within the genus *Mycobacterium.* Int. J. Syst. Bacteriol. 40:323-330.
47. Rozas, J., Sanchez-DelBarrio, J. C., Messeguer, X., and Rozas, R. 2003. DnaSP, DNA polymorphism analysis by the coalescent and other methods. Bioinformatics 19:2496-1497.
48. Rychlik, W. 2002. OLIGO: Primer Analysis Software, version 6.65. Molecular Biology Insights Inc., Cascade, CO.
49. Sarkar, S. F, and Guttman, D. S. 2004. Evolution of the core genome of *Pseudomonas syringae*, a highly clonal, endemic plant pathogen. Appl. Environ. Microbiol. 70:1999-2012.
50. Scally, M., Schuenzel, E. L., Stouthamer, R., and 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.
51. Schaad, N. W., Postnikova, E., Lacy, G., Fatmi, M., and Chang, C. J. 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.
52. Schuenzel, E. L., Scally, M., Stouthamer, R., and 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.
53. Sherald, J. L. 1993. Pathogenicity of *Xylella fastidiosa* in American elm and failure of reciprocal transmission between strains from elm and sycamore. Plant Dis. 77:190-193.
54. Shriver, M. D., Jin, L., Boerwinkle, E., Deka, R., Ferrell, R. E., and Chakraborty, R. 1995. A novel measure of genetic distance for highly polymorphic tandem repeat loci. Mol. Biol. Evol. 12:914-920.
55. Silva, V. S., Shida, C. S., Rodrigues, F. B., Ribeiro, D. C. D., de Souza, A. A., Coletta-Filho, H. D., Machado, M. A., Nunes, L. R., and de Oliveira, R. D. 2007. Comparative genomic characterization of citrus-associated *Xylella fastidiosa* strains. BMC Genomics 8:474.
56. Simpson, A. J. G., Reinach, F. C., Arruda, P., Abreu, F. A., Acencio, M., Alvarenga, R., Alves, L. M. C., Araya, J. E., Baia, G. S., Baptista, C. S., Barros, M. H., Bonaccorsi, E. D., Bordin, S., Bove, J. M., Briones, M. R. S., Bueno, M. R. P., Camargo, A. A., Camargo, L. E. A., Carraro, D. M., Carrer, H., Colauto, N. B., Colombo, C., Costa, F. F., Costa, M. C. R., Costa-Neto, C. M., Coutinho, L. L., Cristofani, M., Dias-Neto, E., Docena, C., El-Dorry, H., Facincani, A. P., Ferreira, A. J. S., Ferreira, V. C. A., Ferro, J. A., Fraga, J. S., Franc, S. C., Franco, M. C., Frohme, M., Furlan, L. R., Garnier, M., Goldman, G. H., Goldman, M. H. S., Gomes, S. L., Gruber, A., Ho, P. L., Hoheisel, J. D., Junqueira, M. L., Kemper, E. L., Kitajima, J. P., Krieger, J. E., Kuramae, E. E., Laigret, F., Lambais, M. R., Leite, L. C. C., Lemos, E. G. M., Lemos, M. V. F., Lopes, S. A., Lopes, C. R., Machado, J. A., Machado, M. A., Madeira, A. M. B. N., Madeira, H. M. F., Marino, C. L., Marques, M. V., Martins, E. A. L., Martins, E. M. F., Matsukuma, A. Y., Menck, C. F. M., Miracca, E. C., Miyaki, C. Y., Monteiro-Vitorello, C. B., Moon, D. H., Nagai, M. A., Nascimento, A. L. T. O., Netto, L. E. S., Nhani, A., Nobrega, F. G., Nunes, L. R., Oliveira, M. A., de Oliveira, M. C., de Oliveira, R. C., Palmieri, D. A., Paris, A., Peixoto, B. R., Pereira, G. A. G., Pereira, H. A., Pesquero, J. B., Quaggio, R. B., Roberto, P. G., Rodrigues, V., de M.Rosa, A. J., de Rosa, V. E., de Sa´, R. G., Santelli, R. V., Sawasaki, H. E., da Silva, A. C. R., da Silva, A. M., da Silva, F. R., Silva, W. A., da Silveira, J. F., Silvestri, M. L. Z., Siqueira, W. J., de Souza, A. A., de Souza, A. P., Terenzi, M. F., Truffi, D., Tsai, S. M., Tsuhako, M. H., Vallada, H., Van Sluys, M. A., Verjovski-Almeida, S., Vettore, A. L., Zago, M. A., Zatz, M., Meidanis, J., and Setubal, J. C. 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. Nature 406:151-159.
57. Spratt, B. G. 1999. Multilocus sequence typing: Molecular typing of bacterial pathogens in an era of rapid DNA sequencing and the Internet. Curr. Microbiol. 2:312-316.
58. Spratt, B. G., Hanage, W. P., and Feil, E. J. 2001. The relative contributions of recombination and point mutation to the diversification of bacterial clones. Curr. Microbiol. 4:602-606.
59. Swofford, D. L. 2002. PAUP*. Phylogenetic Analysis Using Parsimony (* and other methods), version 4. Sinauer Associates, Sunderland, MA.
60. Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585-595.
61. Urwin, R., and Maiden, M. C. J. 2003. Multi-locus sequence typing: A tool for global epidemiology. Trends Microbiol. 11:479-487.
62. Van Sluys, M. A., de Oliveira, M. C., Monteiro-Vitorello, C. B., Miyaki, C. Y., Furlan, L. R., Camargo, L. E. A., da Silva, A. C. R., Moon, D. H., Takita, M. A., Lemos, E. G. M., Machado, M. A., Ferro, M. I. T., da Silva, F. R., Goldman, M. H. S., Goldman, G. H., Lemos, M. V. F., El-Dorry, H., Tsai, S. M., Carrer, H., Carraro, D. M., de Oliveira, R. C., Nunes, L. R., Siqueira, W. J., Coutinho, L. L., Kimura, E. T., Ferro, E. S., Harakava, R., Kuramae, E. E., Marino, C. L., Giglioti, E., Abreu, I. L., Alves, L. M. C., do Amaral, A. M., Baia, G. S., Blanco, S. R., Brito, M. S., Cannavan, F. S., Celestino, A. V., da Cunha, A. F., Fenille, R. C., Ferro, J. A., Formighieri, E. F., Kishi, L. T., Leoni, S. G., Oliveira, A. R., Rosa, V. E., Sassaki, F. T., Sena, J. A. D., de Souza, A. A., Truffi, D., Tsukumo, F., Yanai, G. M., Zaros, L. G., Civerolo, E. L., Simpson, A. J. G., Almeida, N. F., Setubal, J. C., and Kitajima, J. P. 2003. Comparative analyses of the complete genome sequences of Pierce's disease and citrus variegated chlorosis strains of *Xylella fastidiosa*. J. Bacteriol. 185:1018-1026.
63. Vickerman, D., Luck, R. F., Stouthamer, R., Nunney, L., and Cooksey, D. 2005. Multilocus sequence typing (MLST) to identify reservoirs of *Xylella fastidiosa* diversity in natural hosts in California. Pages 181-184 in: Proc. 2005 Pierce's Disease Res. Symp. California Department of Food and Agriculture Pierce's Disease Control Program.
64. Wang, P. W., Morgan, R. L., Scortichini, M, and Guttman, D. S. 2007. Convergent evolution of phytopathogenic pseudomonads onto hazelnut. Microbiology 153:2067-2073.
65. Wells, J. M., Raju, B. C., Hung, H. Y., Weisburg, W. G., Mandelco-Paul, L., and Brenner, D. J. 1987. *Xylella fastidiosa*: Gram-negative, xylem-limited, fastidious plant bacteria related to *Xanthomonas*. Int. J. Syst. Bacteriol. 37:136-143.
66. Wichman, R. L., and Hopkins, D. L. 2002. Differentiation of pathogenic groups of *Xylella fastidiosa* strains with whole-cell protein profiles. Plant Dis. 86:875-879.
67. Yan, S., Liu, H., Mohr, T. J., Jenrette, J., Chiodini, R., Zaccardelli, M., Setubal, J. C., and Vinatzer, B. A. 2008. Role of recombination in the evolution of the model plant pathogen *Pseudomonas syringae* pv. tomato DC3000, a very atypical tomato strain. Appl. Environ. Microbiol. 74:3171-3181.
68. Yang, Z. 1998. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. Mol. Biol. Evol. 15:568-573.