

# Evaluation of the Genetic Diversity of *Xylella fastidiosa* Strains from Citrus and Coffee Hosts by Single-Nucleotide Polymorphism Markers

Ester Wickert, Marcos Antonio Machado, and Eliana G. M. Lemos

First and third authors: Microorganism and Plant Biochemistry Laboratory, Department of Technology, School of Agricultural and Veterinary Sciences, University of the State of São Paulo (UNESP), Jaboticabal, SP, Brazil; and second author: Biotechnology Laboratory, APTA Citrus/IAC, State of São Paulo, Cordeirópolis, SP, Brazil.

Accepted for publication 15 June 2007.

## ABSTRACT

Wickert, E., Machado, M. A., and Lemos, E. G. M. 2007. Evaluation of genetic diversity of *Xylella fastidiosa* strains from citrus and coffee hosts by single-nucleotide polymorphism markers. *Phytopathology* 97:1543-1549.

The aim of this study was to obtain information about genetic diversity and make some inferences about the relationship of 27 strains of *Xylella fastidiosa* from different hosts and distinct geographical areas. Single-nucleotide polymorphism (SNP) molecular markers were identified in DNA sequences from 16 distinct regions of the genome of 24 strains of *X. fastidiosa* from coffee and citrus plants. Among the Brazilian strains,

coffee-dependent strains have a greater number of SNPs (10 to 24 SNPs) than the citrus-based strains (2 to 12 SNPs); all the strains were compared with the sequenced strain 9a5c. The identified SNP markers were able to distinguish, for the first time, strains from citrus plants and coffee and showed that strains from coffee present higher genetic diversity than the others. These markers also have proven to be efficient for discriminating strains from the same host obtained from different geographic regions. *X. fastidiosa*, the causal agent of citrus variegated chlorosis, possesses genetic diversity, and the SNP markers were highly efficient for discriminating genetically close organisms.

*Xylella fastidiosa*, originally characterized by Wells et al. (50), is epidemic in Brazil, where it causes a disease known as “clorose variegada dos citros” (citrus variegated chlorosis [CVC]), which is responsible for great losses in a major crop, citrus. This pathogen was the first phytopathogenic plant pathogen to have its genome completely sequenced (47). It also affects coffee, another economically important crop in Brazil, causing coffee leaf scorch (CLS) (28,36,39).

Strains from *X. fastidiosa* also are related to other diseases that affect various crops in many countries, such as Pierce’s disease in grapevines, alfalfa dwarf, plum leaf scald, and phony peach. There have been reports of related diseases affecting pear and pecan trees (18,26,43). Strains of *X. fastidiosa* interact with different hosts (18,19,46), have different nutritional requirements (12,18,20), possess different degrees of pathogenicity (12,17,21), and have different DNA homology (3,4,22). Recently, three *Xylella* subspecies were described (44) based on a DNA-DNA relatedness assay and comparison of the 16S-23S rDNA intergenic sequences. These authors proposed *X. fastidiosa* subsp. *piercei*, *X. fastidiosa* subsp. *multiplex*, and *X. fastidiosa* subsp. *pauca* as new subspecies.

Methods to distinguish *X. fastidiosa* strains are fundamental for ecological and epidemiological studies, especially due to its widespread occurrence. Studies of genetic diversity of *X. fastidiosa* from different hosts have used different molecular techniques such as repetitive extragenic palindromic polymerase chain reaction (rep-PCR), random amplified polymorphic DNA (RAPD), restriction fragment length polymorphism (RFLP), PCR-RFLP, variable number of tandem repeats (VNTR), and others, such as pulsed field gel electrophoresis (PFGE) (3,4,6,8,9,16,37–40).

Although molecular techniques were able to determine the existence of genetic variability among strains from different hosts, strains obtained from citrus plants and coffee could not be differentiated by most of the markers previously used.

Single-nucleotide polymorphism (SNP) markers correspond to base alterations (mutations) associated with deletions, insertions, or substitutions in the DNA (10,15,49). Because they correspond to alterations of a single base, they are potential markers for studies with genetically closely related organisms.

Although genetic diversity among *X. fastidiosa* strains from different hosts already has been detected, there is little evidence of genetic diversity of strains from the same host. Recently, using SNPs in the 16S rDNA, two genotypes of *X. fastidiosa* were found associated with almond leaf scorch in the same location in California (5). Moreover, there is little knowledge about the phylogenetic relationships between strains obtained from the same host and those obtained from different hosts, especially for citrus plants and coffee. These two crops are economically very important for Brazil and are cultivated in subtropical areas and, sometimes, orchards occur close to each other, a fact that could facilitate dissemination of vectors and, consequently, the bacterium *X. fastidiosa*.

Because citrus and coffee crops coexist temporally and geographically, a determination of the level of genetic subdivision of *X. fastidiosa* strains from both crops will guide important decisions regarding how to deal with this bacterium. This knowledge will be useful for the control of diseases such as CVC and CLS. It is also important to determine whether geographic isolation of foreign strains from grapevines has an effect on the phylogenetic relationships of native strains from coffee and citrus plants.

In this study, 16 arbitrarily selected DNA sites identified in the strain 9a5c genome were sequenced in 26 *X. fastidiosa* strains from grapevine, citrus, and coffee crops. DNA sequences in these 16 regions were used to verify genetic diversity among strains from the same host and to infer the phylogenetic relationship of these strains.

Corresponding author: E. G. M. Lemos; E-mail address: egerle@fcav.unesp.br

doi:10.1094/PHYTO-97-12-1543

© 2007 The American Phytopathological Society

## MATERIALS AND METHODS

**Strains and culture conditions.** Twenty-six strains from different hosts were kindly provided by the Biotechnology Laboratory at the Sylvio Moreira Citriculture Center (CCSM/IAC), the Agronomy Institute of the State of Paraná (IAPAR), and the Microorganism and Plant Biochemistry Laboratory (LBMP) in the Technology Department, FCAV/UNESP (School of Veterinary and Agricultural Sciences of the São Paulo State University) (Table 1). Strains were grown on XDM<sub>2</sub> medium (25).

**DNA extraction.** Extraction of DNA was carried out according to the protocol described by Ausubel (2). DNA was quantified using a BECKMAN-DU 640 spectrophotometer (42).

**Design of region-specific primers.** Specific primers were designed for the 16 DNA sites that were selected arbitrarily from the *X. fastidiosa* genome project database, such that the amplified products exhibited a size of ≈400 bp. Forward and reverse primers were designed using Gene Runner software. The 16 selected regions are listed in Table 2. The primers, with their respective sequence forward (5' to 3') and reverse (3' to 5'), the melting temperature (T<sub>m</sub>) for each primer, and the T<sub>m</sub> used for amplification with the primer pair are listed in Table 2.

**Confirmation of *X. fastidiosa* strains.** In order to identify whether the strains belong to the *X. fastidiosa* species, a PCR reaction was performed using primers RST31 and RST33 (32), which are specific primers for this bacterium. PCR was carried out in a 20-μl reaction containing 2 μl of 1× buffer (50 mM KCl and 200 mM Tris-HCl, pH 8.4), 0.8 μl of 5 mM MgCl<sub>2</sub>, 0.4 μl of 10 mM each dNTP, 0.2 μl of Taq DNA polymerase, 5 pmol of each primer, and 50 ng of genomic DNA, and was brought up to volume using pure sterile water. Reactions were performed in a thermal cycler (PTC-100 Programmable Thermal Controller; MJ Research, Inc.) using an initial cycle of 94°C for 2 min; 39 cycles of 94°C for 1 min, 1 min at 62°C, and 72°C for 90 s; and a final cycle of 72°C for 5 min. Amplified samples were electrophoresed

in 1.5% agarose gel containing ethidium bromide (0.5 μg/ml). Samples then were observed under UV light and photographed using a documentation system (GEL DOC 1000; BioRad).

**Amplification, purification, and sequencing of the regions.** Primers were used for target amplification of the 27 strains, in PCR containing 100 ng of DNA from each strain in a total volume of 100 μl as follows 10 μl 1× buffer (50 mM KCl, 200 mM Tris-HCl, pH 8.4), 4 μl of 5 mM MgCl<sub>2</sub>, 2 μl of 10 mM dNTP, 1 μl of Taq DNA polymerase, 5 pmol of each primer, and sterile water to bring the reaction to 100 μl. Reactions were performed in a thermal cycler (PTC-100 Programmable Thermal Controller; MJ Research, Inc.), altering the T<sub>m</sub> to the recommended temperature for each pair of primers (Table 2). The other parameters of the amplification cycles were as previously described. PCR products were purified using QIAquick 96 PCR Purification Kit, according to manufacturer's instructions (Qiagen, Germany). Purified DNA was used for PCR using the DNA Sequencing-Big Dye Terminator Cycle Sequencing Ready Reaction ABI Prism Kit (2 μl of Big Dye terminators, 5 pmol of forward and reverse primers, and 50 ng of DNA in 8 μl of pure sterile water). Reactions again were carried out in a thermal cycler (PTC-100 Programmable Thermal Controller; MJ Research, Inc.) using the following cycles: 2 min at 94°C; 35 cycles of 94°C for 1 min, T<sub>m</sub> in °C for 1 min; and 72°C for 1 min and 30 s; and 5 min at 72°C. Samples containing fragments were precipitated with 1 ml of isopropanol (75%), washed with 1 ml of ethanol (70%), resuspended in 3 μl of standard loading buffer (5:1 deionized formamide/50 mM EDTA, pH 8.0) at 95°C for 2 min, and then applied to a sequencing gel in an ABI377 sequencer (Perkin Elmer). Amplified regions were sequenced three times to confirm the DNA sequence obtained.

**Sequence analysis.** Sequences were analyzed using the Phred/Phrap/Consed program package (13) to verify the quality of DNA sequence obtained. Sequences were aligned using the Sequencher program, and LINSEQ (sequence linerizer program) was used for concatenation of all regions. The 16 sequences from each strain

TABLE 1. Bacterial strains used in this study and their specifications<sup>a</sup>

Strains	Number of SNPs	Frequency of SNPs	Host	Geographic origin <sup>b</sup>	Source <sup>c</sup>
9a5c	0	0	<i>Citrus sinensis</i>	Macaubal, SP	LBMP
U150a	5	0.0010488	<i>C. sinensis</i>	Ubarana, SP	CCSM
U161b	2	0.0004195	<i>C. sinensis</i>	Ubarana, SP	CCSM
B16	8	0.0016778	<i>C. sinensis</i>	Bebedouro, SP	CCSM
GP920b	8	0.0016778	<i>C. sinensis</i>	Gavião Peixoto, SP	CCSM
ITA B	8	0.0016778	<i>C. sinensis</i>	Itapetininga, SP	CCSM
SR1316A	7	0.0014684	<i>C. sinensis</i>	Santa Rita do Passaquatro, SP	CCSM
10438	7	0.0014684	<i>C. sinensis</i>	Loanda, PR	IAPAR
11038	11	0.0023075	<i>C. sinensis</i>	São Jorge, SP	IAPAR
11066	7	0.0014684	<i>C. sinensis</i>	Londrina, PR	IAPAR
11067	6	0.0012586	<i>C. sinensis</i>	Londrina, PR	IAPAR
11347	7	0.0014684	<i>C. sinensis</i>	Loanda, PR	IAPAR
11348	7	0.0014684	<i>C. sinensis</i>	São Pedro do Paraná, PR	IAPAR
11380	7	0.0014684	<i>C. sinensis</i>	Pinhalzinho, SC	IAPAR
11400	12	0.0025173	<i>C. sinensis</i>	Boquim, SE	IAPAR
11779	10	0.0020977	<i>C. sinensis</i>	Pinhalzinho, SC	IAPAR
12288	9	0.0018879	<i>C. sinensis</i>	Londrina, PR	IAPAR
C3	12	0.0025173	<i>Coffea arabica</i>	Cordeirópolis, SP	CCSM
C6	13	0.0027270	<i>C. arabica</i>	Cordeirópolis, SP	CCSM
C10	10	0.0020977	<i>C. arabica</i>	Cordeirópolis, SP	CCSM
C11	13	0.0027270	<i>C. arabica</i>	Cordeirópolis, SP	CCSM
JAB1	17	0.0035666	<i>C. arabica</i>	Garça, SP	LBMP
JAB2	24	0.0050346	<i>C. arabica</i>	Garça, SP	LBMP
JAB14	15	0.0031466	<i>C. arabica</i>	Garça, SP	LBMP
8935	124	0.0260121	<i>Vitis vinifera</i>	Florida, USA	ATCC 35879
9713	135	0.0283196	<i>V. vinifera</i>	Florida, USA	IAPAR
9715	127	0.0266414	<i>V. vinifera</i>	Florida, USA	IAPAR

<sup>a</sup> SNP = single-nucleotide polymorphism.

<sup>b</sup> SP = state of São Paulo, Brazil, PR = state of Paraná, Brazil, SC = state of Santa Catarina, Brazil, SE = state of Sergipe State, Brazil, and USA = United States of America.

<sup>c</sup> LBMP = Plants and Microorganism Biochemistry Laboratory at Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal, Universidade Estadual de São Paulo, São Paulo, Brazil; CCSM = Sylvio Moreira Citriculture Center, Centro Apta de Citros, Instituto Agronômico, Cordeirópolis, São Paulo, Brazil; IAPAR = Instituto Agronômico do Paraná, Paraná, Brazil; and ATCC = American Type Culture Collection, Manassas, VA.

were placed in consecutive order according to their position in the genome of reference strain 9a5c and linked to form a unique sequence for each strain. Strain 9a5c was used as a reference (47).

**Sequence data submission.** Sequences used in this study were submitted to EMBL. Sequences were named according to strain number and the amplified gene, and the respective accession numbers are in Table 3.

**Genetic diversity.** Genetic diversity indices were calculated using the DNA sequences of the strains from coffee and citrus plants. Grapevine strains, which probably belong to the *X. fastidiosa* subsp. *piercei*, as described previously by Schaad et al. (44), were not used to calculate diversity indices because they are not from Brazil. They were used only as an outgroup to build a dendrogram to analyze the genetic relationships among the Brazilian strains. Strains from citrus plants and coffee were placed in groups according to their respective original hosts for the genetic diversity analysis.

**Average SNP frequency.** Values of average SNP frequency were calculated ( $f \text{ SNP} = \text{total SNP number} / \text{total number of bases in the sequence} / \text{number of strains in the sample}$ ), and results are

listed in Table 1 for each strain; the average for strains from the same host are presented in Table 4.

**Genetic distance.** The genetic distance was calculated between the strains from citrus and coffee crops (Table 4). Estimates of genetic distances were used to evaluate genetic divergence within and between groups of strains (34). The genetic distance within the citrus- and coffee-based groups (infrasubspecific) was estimated by the arithmetic mean of all individual pairwise distances between taxa within a group. The genetic distance between (inter-subspecific) citrus-, coffee-, and grape-based groups was estimated by the arithmetic mean of all pairwise distances between the two groups in the intergroup comparisons (35). These values were calculated using the Kimura-2-parameter nucleotide substitution method (23) using MEGA software (2.1 version) (24).

**Other infrasubspecific genetic diversity indices.** Values of infrasubspecific genetic diversity were calculated to estimate other genetic diversity indices between the groups of strains from citrus plants and coffee using Arlequin software (45). Average pairwise differences were estimated from the intragroup comparisons of the number of sequence differences between a given strain and all

TABLE 2. List of primers, with their respective pairing regions in the *Xylella fastidiosa* genome, their forward (5' to 3') and reverse (3' to 5') sequences, their respective melting temperature (Tm) (°C), and the Tm (°C) used for amplification with the primer pair

Primers	Sequence 5' to 3' - Tm (°C)	Sequence 3' to 5' - Tm (°C)	Genes	No. of SNPs <sup>a</sup>	Used Tm (°C) <sup>b</sup>
22/23	TGGAGGGTGGGTGATACTTGGG - 70	CGCGGCAAAGCATGTTCATATTG - 66	<i>XF 2111 - untranslated sequence</i>	13	66
24/25	TTGACGGTAGGCGGTATGTGGC - 70	GCCTGCATGAAGTCCAAACCAG - 68	<i>XF 0919 - conserved hypothetical protein</i>	4	66
44/45	ACACCTTTCTCCCTGATTGCG - 64	GGAACGGGTAGCAGACTACGG - 70	<i>XF 2111 - hypothetical protein</i>	11	64
46/47	TACTGGCGGCTCTCCCTTC - 64	CCTGGTAAGGGATTGTGTGACG - 68	<i>XF 0595 - hypothetical protein</i>	17	64
48/49	TTCTGCTCGCACCTAATCC - 66	GCTTGGGTGTCTGAGCAGGTG - 70	<i>XF 2364 - Gum H protein (hexapolsaccharidae)</i>	6	66
62/63	GCTCCCGAACCTCGTAATC - 64	GATTTGCGTGCTGGAACGTTAG - 66	<i>XF 0752 - conserved hypothetical protein</i>	11	64
68/69	GACCATCTCAAGCAGCAACTC - 68	AGCGCAGTGCATTATCCAGG - 62	<i>XF 1363 - soluble lytic murein transglycosylase precursor</i>	6	62
74/75	AGCACCACCACAACCAGTAACC - 68	TCGGAGCGTTGGAGCTTTC - 60	<i>XF 0103 - membrane protein</i>	16	64
80/81	GCCAAACAAGCACGGGACAAAG - 68	TGCGGGTTCATTTCTCCTAGGC - 68	<i>XF1286 - topoisomerase IV subunit B</i>	19	66
82/83	CGATTTGGTGCTGACTGATGTG - 66	GCAGGAAGGCAGACAAATCCC - 66	<i>XF 2693 - hypothetical protein</i>	3	66
92/93	CCCCTGCCTCATTCACTGC - 64	CCAACAAGGCAGCAACAACAG - 64	<i>XF 1640 - ankyrin-like protein</i>	5	64
96/97	GCAGAACAGCGGTGGCAGATAAG - 72	CGTCATTGCACGACCACACTC - 66	<i>XF 0144 - biosynthetic arginine decarboxylase</i>	15	66
104/105	TGAAGGTGTGGGCTGGAATGCG - 70	GGGTTTCGACGTTGAGTGAAGC - 66	<i>XF 0749 - untranslated sequence</i>	29	66
106/107	GTCTCCAACCGTCCGAAAC - 64	GGTTCGCTACCAAGTGTTCAG - 70	<i>XF 2247 - hypothetical protein</i>	2	66
108/109	CCAGCGCATAGCGTTGTGAAG - 70	TTAGCCGGTTTCAGCTTTGGAG - 66	<i>XF 1829 - conserved hypothetical protein</i>	11	66
110/111	GCCGCCAAGCAATCAACGAC - 64	GARACTCATGCGGCAGCACC - 62	<i>XF 2133 - ABC transporter-binding protein</i>	15	63

<sup>a</sup> SNPs = single-nucleotide polymorphisms.

<sup>b</sup> Tm used for amplification with the primer pair.

TABLE 3. Accession numbers for the DNA sequences from the 16 different regions of the 27 *Xylella fastidiosa* strains deposited in GenBank

Strains <sup>a</sup>	Genes	Accession number
1, 2, 3, 4-27	XF 0103	DQ370506; DQ370515; DQ370517; DQ370524-DQ370547
1, 2, 3, 4-27	XF 0144	DQ370548; DQ370557; DQ370559; DQ370566-DQ370589
1, 2, 3, 4-27	XF 0595	DQ370590; DQ370599; DQ370601; DQ370608-DQ370631
1, 2, 3, 4-27	XF 0749	DQ370632; DQ370641; DQ370643; DQ370650-DQ370673
1, 2, 3, 4-27	XF 0752	DQ370674; DQ370683; DQ370685; DQ370692-DQ370715
1, 2, 3, 4-27	XF 0919	DQ370716; DQ370725; DQ370727; DQ370734-DQ370757
1, 2, 3, 4-27	XF 1286	DQ370758; DQ370767; DQ370769; DQ370776-DQ370799
1, 2, 3, 4-27	XF 1363	DQ370800; DQ370809; DQ370811; DQ370817-DQ370841
1, 2, 3, 4-27	XF 1640	DQ370842; DQ370851; DQ370853; DQ370860-DQ370883
1, 2, 3, 4-27	XF 1829	DQ370884; DQ370893; DQ370895; DQ370902-DQ370925
1, 2, 3, 4-27	XF 2111	DQ370926; DQ370935; DQ370937; DQ370944-DQ370967
1, 2, 3, 4-27	XF 2111a	DQ370968; DQ370977; DQ370979; DQ370986-DQ371009
1, 2, 3, 4-27	XF 2133	DQ371010; DQ371019; DQ371020; DQ371028-DQ371051
1, 2, 3, 4-27	XF 2247	DQ371052; DQ371061; DQ371063; DQ371070-DQ371093
1, 2, 3, 4-27	XF 2364	DQ371094; DQ371103; DQ371105; DQ371112-DQ371135
1, 2, 3, 4-27	XF 2693	DQ371136; DQ371145; DQ371147; DQ371154-DQ371177

<sup>a</sup> Strain numbers = 1: 9a5c; 2: U150a; 3: U161b; 4: B16; 5: GP920b; 6: ITA B; 7: SR1316A; 8: 10438; 9: 11038; 10: 11066; 11: 11067; 12: 11347; 13: 11348; 14: 11380; 15: 11400; 16: 11779; 17: 12288; 18: C3; 19: C6; 20: C10; 21: C11; 22: JAB1; 23: JAB2; 24: JAB14; 25: 8935; 26: 9713; 27: 9715.

other strains (45) (Table 4). To estimate genetic diversity within the groups of strains from citrus plants and coffee, the indices were calculated using the distance method with Kimura-2-parameter substitution nucleotide model. Average pairwise differences and nucleotide diversity were calculated for each group. Also, molecular indices, such as the number of used sequences (gene copies) and haplotypes, total number of usable loci, polymorphic sites, and nucleotide diversity, were calculated for each data set (Table 4). Nucleotide diversity was estimated from the number of variable positions for the aligned sequences in a given DNA sequence group. The same software was used to calculate the shared haplotypes by citrus and coffee hosts.

**Phylogeny.** Aligned and concatenated sequences were used to build a phylogenetic tree to infer the relationships among all strains. These relationships were inferred by the distance method and algorithm Neighbor Joining (41) with Kimura-2-parameter nucleotide substitution model (23) using MEGA software with a 1,000-repetition bootstrap (11). Grapevine strains were used as the outgroup.

## RESULTS

**Primer design and confirmation of *X. fastidiosa* strains.** It was possible to design several primers, but only the 16 selected pairs successfully amplified the fragments. All strains used in this work were submitted to amplification with specific primers RST31 and RST33 (32) and showed the 733- or 735-bp product expected for *X. fastidiosa* strains (733 bp for Pierce's disease strains and 735 for CVC strains) (data not shown).

**Sequence analysis.** The Phred/Phrap/Consed program package was used to verify the quality of obtained sequences. Because single base alterations are used to differentiate the strains, this high quality standard is absolutely necessary; any problems regarding quality of the sequences could negatively affect the accuracy of final result. A total of 4,767 bp of each strain was used for this analysis.

**Genetic diversity.** Several indices were calculated in order to estimate the genetic diversity between and among citrus- and coffee-derived strains. When the DNA sequence is used as a source of information, diversity is defined as the nucleotide sequence variation between individuals, assuming all sample sequences are related to a common ancestor (29).

**Average SNP frequency.** Different numbers of SNPs were found in the 16 DNA regions (Table 2) among strains from different hosts. Strain 9a5c (47) was used as the standard, with SNP frequency defined as 0 (Table 1). SNP frequencies also were calculated for each different host strain. *X. fastidiosa* strains obtained from grapevines presented the highest SNP frequency, followed by coffee strains and then citrus strains. Among the citrus strains, the highest SNP frequency was observed in strain 11400, with a value of 0.0025173 (12 SNPs). All other strains presented SNP frequency values between 0.0004195 (2 SNPs) and 0.0023075 (11 SNPs). Regarding the strains isolated from coffee plants, the highest SNP frequency was observed in strain JAB 2: 0.0050346

(24 SNPs), and the lowest in strain C10: 0.0020977 (10 SNPs). Grapevine strains presented significantly higher SNP values compared with citrus and coffee. The highest number and frequency of SNP in grapevines were found in strain 9713 (135 and 0.02831), respectively.

**Genetic distance.** Genetic distance indices were estimated for citrus and coffee strains (Table 4). Genetic distance indices were calculated to evaluate genetic divergence within and between coffee and citrus strains. Low values of genetic distance indicate that the organisms or groups are genetically close, and higher distance values indicate greater genetic divergence. The greatest genetic distance among strains from the same host was observed for coffee (0.003984), followed by citrus (0.001731). The greater genetic distance found between coffee plant strains shows the higher genetic diversity of this group compared with citrus, despite the fact that coffee strains are present in lower numbers than citrus strains. The genetic distance between citrus and coffee strains presents a low value (0.004021) compared with the genetic distance from grapevine strains in comparison with coffee and citrus strains (0.02665 and 0.02710, respectively). This indicates that Brazilian coffee and citrus strains are genetically close.

**Other infrasubspecific genetic diversity indices.** The greatest average pairwise difference values were presented by coffee strains as a group and these strains also presented higher values for the number of polymorphic sites and nucleotide diversity than did the citrus strains. All measures of genetic diversity agree that coffee strains as a group have higher variability than the citrus strains. No shared haplotypes were found among strains from citrus and coffee plants (Table 4).

**Phylogeny.** An inference of phylogenetic relationships for the strains is shown in Figure 1. Strains used in this study form groups according to their hosts and geographical origin. The phylogenetic tree was built using the 16 linearized DNA sequences from 24 strains from coffee and citrus plants and 3 sequences from grapevine strains, with the objective of demonstrating the genetic relationships among them. Five large groups of strains could be observed: the branch formed by the grapevine strains from the United States, which were used as the outgroup; a group formed by coffee strains JAB1, JAB2, and JAB14, native to Garça, Sao Paulo State, and C11, from Cordeiropolis, São Paulo State (coffee A); a group of coffee plant strains C3, C6, and C10 from Cordeiropolis, Sao Paulo State (coffee B); a branch formed by citrus strains 11779, 11067, 10438, 11347, 11380, 11348, and 11066 from Santa Catarina and Paraná States (citrus A); and a branch from citrus strains 9a5c, U161b, U150a, GP920b, 11038, ITA B, B16, SR1316A, 12288, and 11400 from Sao Paulo and Sergipe States, which was named citrus B.

## DISCUSSION

In this study, 24 *X. fastidiosa* strains from citrus (17) and coffee (7) crops were used to analyze the genetic diversity among same host strains (infrasubspecific) and different host strains (intersubspecific). Three additional strains from grapevines (*X. fastidi-*

TABLE 4. Values of several diversity indices calculated for strains from citrus and coffee crops

Index	Citrus-derived	Coffee-derived
Average SNP frequency <sup>a</sup>	0.00119	0.00312
Infrasubspecific genetic distance	0.00173	0.00398
Number of sequences used	17	7
Number of haplotypes	15	7
Total number of sites	4,767	4,767
Number of polymorphic sites	30	55
Nucleotide diversity	0.0019 ± 0.0010	0.0047 ± 0.0027
Average pairwise difference	9.3676 ± 4.5268	23.0000 ± 11.5603
Number of haplotypes shared by hosts		0
Intersubspecific genetic distance		0.0042

<sup>a</sup> SNP = single-nucleotide polymorphism.

osa subsp. piercei) were used as outlying strains to build a phylogenetic tree for the strains used in this study. The DNA sequence data used was obtained from 16 arbitrarily distributed regions in the bacterial genome.

Among Brazilian strains, coffee strains presented a higher frequency of SNPs and infrasubspecific genetic distances than citrus. These values of infrasubspecific genetic distances data also showed greater genetic diversity among coffee-based strains than in citrus. Citrus values are in agreement with previously reported genetic diversity data obtained for strains from clonal citrus plants and for strains obtained from different orange cultivars (7). Coffee strains have greater nucleotide diversity, a larger number of polymorphic sites, and more pairwise differences than citrus strains.

Therefore, the number of different *X. fastidiosa* strains associated with coffee probably is greater than those associated with citrus. Coffee (*Coffea arabica*) is a self-pollinating plant with a high percentage of open pollinations and has been propagated by seed for many years in Brazil. Because these orchards were not planted with clonal plants, there was genetic diversity among plants from the same orchard. In Brazil, coffee is an important crop, widely cultivated and spread in subtropical Brazilian areas since 1790. The long period of coffee production in Brazil, allied with genetic diversity among plants from the same farm, likely have influenced the genetic structure of pathogens like *X. fastidiosa*. This might explain the greater genetic diversity among coffee-based strains when compared with citrus-based strains. The host probably exerts selection pressure on the bacterium, causing it to adapt following inoculation by selecting for DNA alterations and causing the microorganism to adapt to this particular host—one plant from the orchard—leading to genetic diversity. On the other hand, Brazilian citrus orchards are more recent, since ≈1930, and cultivated varieties are propagated vegetatively. The orchards are essentially clonal, the culture is recent, and, therefore, pathogens are under less selection pressure and do not have great need for diversification.

The fact that *X. fastidiosa* strains from coffee and citrus are genetically related suggests that they may have a common ancestor. The great genetic diversity of coffee strains and the fact that coffee-based strains are genetically close to citrus-based strains—when compared with grapevine-derived strains—suggests that this microorganism may have migrated from coffee plants to citrus, as previously reported (7,31,36,39). Moreover, it is probable that populations with greater infrasubspecific genetic diversity are more adaptable to different conditions and environments and are more likely to leave viable descendants if the environment or potential plant hosts change. This factor can increase their ability to successfully colonize different hosts. This can justify the hypothesis that strains from coffee have migrated to citrus plants. In Brazil, this hypothesis is reinforced by the fact that coffee and citrus crops coexist in widespread regions—citrus orchards are cultivated in subtropical areas along with coffee—such as São Paulo State. This can be favorable to microorganism transmission from one crop to another by insect vectors.

However, *X. fastidiosa* was detected in coffee plants only years after its identification in citrus plants, when the bacteria was detected in coffee and was related to CLS (28). Despite the fact that *X. fastidiosa* was detected in citrus before coffee, it likely had colonized coffee orchards for a much longer time, and its symptoms were confounded initially with nutritional deficiencies and root diseases (28).

Basically, genetic diversity is generated and maintained by ecological and genetic factors. The opportunity to colonize the host and competition among strains are among the ecological factors. Genetic factors include mutations and recombinations (48). In microorganism populations, the main factors responsible for genetic diversity and population structure are recombination and punctual mutations (30). The factors responsible for the genetic diversity seen in *X. fastidiosa* strains from citrus and coffee plants

possibly include climate conditions, population dynamics of the different vectors, and the different types of plants that can be natural hosts for the bacteria. Regarding genetic factors, this study demonstrates the presence of SNPs. The fact that this organism is known to contain phages, transposons, and plasmids (16,47), which are known to be responsible for facilitating the exchange of genetic material between microorganisms (1), favors recombination. In Brazil, infrasubspecific recombination possibly is favored by the presence of more than one haplotype of the microorganism in the same citrus plant, as previously reported (7). A similar phenomenon probably occurs in the coffee plant, but further investigation is necessary in order to verify the presence of one or more *X. fastidiosa* haplotypes in the same coffee plant to confirm this hypothesis.

Ecological factors, such as the process of haplotype selection due to host specificity (14,30), also may be present. Previous studies report that the selection pressure exerted by the host is crucial to determine the genetic structure of a pathogen population (1,14,27,51). The results from the present study show that the haplotypes are probably host-specific because no haplotype was shared by the citrus and coffee plants sampled. However, citrus plants probably can share haplotypes because only 15 haplotypes were found among 17 strains collected from different citrus plants, showing that the same haplotype—at least in these 16 regions—had colonized more than one plant (Table 4). Colleta-Filho and Machado (7) reported similar results when they verified that different sweet orange cultivars were colonized by the same *X. fastidiosa* strains. This was not observed for coffee.

Geographically distinct regions could be one factor that may cause subdivision among microorganism populations. Citrus and coffee strains clearly are distributed according to their hosts as well as to their geographic origin in the phylogram. Great similarity can be observed within coffee and citrus strain groups. Separation of the coffee plant strains into two groups probably reflects their different geographic origin. Regarding the citrus

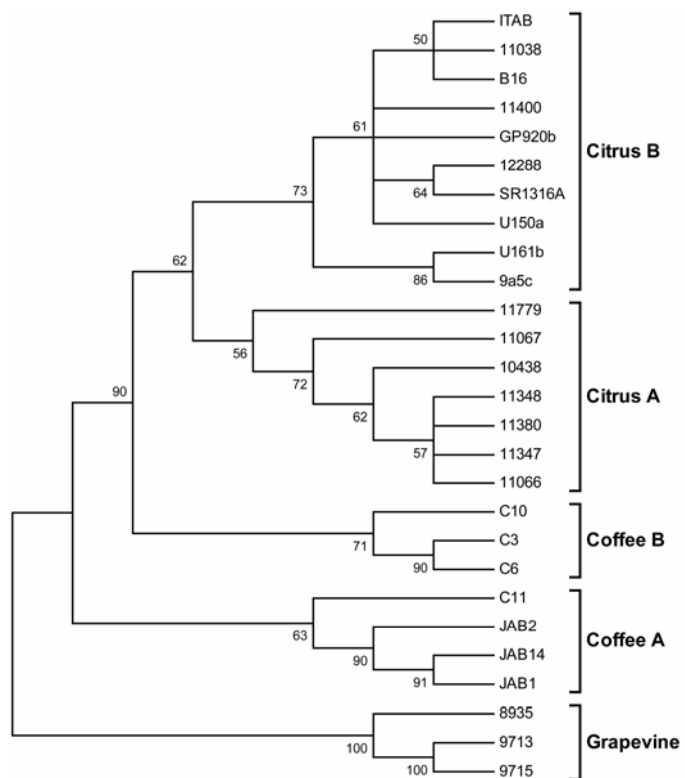


Fig. 1. Relationships among *Xylella fastidiosa* strains from different hosts, showing that they group according to host and geographic origin. Bootstrap values are on the tree.

strains, a subgroup also was formed, possibly reflecting their geographic origin (states of Santa Catarina and Paraná), whereas all other citrus strains came from the states of Sergipe (11400) and São Paulo. A greater similarity between strains from São Paulo and Sergipe has been reported previously (31).

Citrus and coffee strains are genetically closer to each other than grapevine strains, agreeing with the findings of Schaad et al. (44), who classified *Xylella* strains as *X. fastidiosa* subsp. *piercei* (strains from grape, alfalfa, maple, and almond), *X. fastidiosa* subsp. *multiplex* (strains from peach, plum, almond, elm, pigeon grape, sycamore, and other shade trees), and *X. fastidiosa* subsp. *pauca* (citrus strains). These authors did not use coffee strains in their analysis; however, previous studies demonstrated that coffee strains are genetically closer to citrus than to strains from other hosts (39). In this study, citrus and coffee strains could be separated according to their host, but diversity indices demonstrated that they are genetically close. This discrimination could be possible because we used single base modifications as markers in 16 different DNA regions. In this way, Brazilian strains from coffee probably also belong to *X. fastidiosa* subsp. *pauca*, and the observed genetic diversity can be credited to infrasubspecific genetic variation.

This study verified the existence of SNPs among *X. fastidiosa* strains from the same host and from different hosts (coffee and citrus) and the fact that these SNPs may be used to measure infrasubspecific and intersubspecific diversity. This work also inferred the relationships among these strains. We verify that SNPs were able to evaluate genetic diversity among strains from the same host and showed that this diversity is greater among coffee plant strains compared with citrus, and that citrus and coffee strains are genetically close. Considering the hypothesis that strains from coffee plants migrated to citrus and possibly to other hosts, population studies about recombination and gene flow are necessary to obtain information about how these populations are related in space and time.

The bacterium *X. fastidiosa* can be considered a good model for epidemiological studies because the entire genome has been sequenced, it is widespread in the country, and it can colonize different hosts. SNP markers also might become the marker of choice for many applications in population ecology, evolution, and conservation genetics because of the potential for higher genotyping efficiency, data quality, genome-wide coverage, and analytical simplicity (33), helping to infer genetic diversity, population structure, and evolution of plant pathogens. Knowledge about genetic diversity, population structure, and evolution of plant pathogens can help plant breeders make decisions about the use of one or more resistance genes or how to live together with plant diseases and minimize their damage.

## ACKNOWLEDGMENTS

This study was supported by the Foundation for the Support of Research in the State of São Paulo, Brazil (Fundação de Amparo à Pesquisa do Estado de São Paulo–FAPESP). We thank M. V. F. Lemos for reviewing the manuscript.

## LITERATURE CITED

- Arber, W. 2000. Genetic variation: molecular mechanisms and impact on molecular evolution. *FEMS Microbiol. Rev.* 24:1-7.
- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., and Struhl, K. 1988. *Current Protocols in Molecular Biology*. J. Wiley & Sons, New York.
- Chen, J., Chang, C. J., and Jarret, R. L. 1992. Plasmids from *Xylella fastidiosa* strains. *Can. J. Microbiol.* 38:993-995.
- Chen, J., Chang, C. J., Jarret, R. L., and Gawell, N. 1992. Genetic variation among *Xylella fastidiosa* strains. *Phytopathology* 82:973-977.
- Chen, J., Groves, R., Civerolo, E. L., Viveros, M., Freeman, M., and Zheng, Y. 2005. Two *Xylella fastidiosa* genotypes associated with almond leaf scorch disease on the same location in California. *Phytopathology* 95:708-714.
- Chen, J., Lamikanra, O., Chang, C. J., and Hopkins, D. L. 1995. Randomly amplified polymorphic DNA analysis of *Xylella fastidiosa* Pierce's disease and oak leaf scorch pathotypes. *Appl. Environ. Microbiol.* 61:1688-1690.
- Coletta-Filho, H. D., and Machado, M. A. 2002. Evaluation of the genetic structure of *Xylella fastidiosa* populations from different *Citrus sinensis* varieties. *Appl. Environ. Microbiol.* 68:3731-3736.
- Coletta-Filho, H. D., Takita, M. A., Souza, A. A., Aguilar-Vildoso, C., and Machado, M. A. 2001. Differentiation of strains of *Xylella fastidiosa* by a variable number of tandem repeat analysis. *Appl. Environ. Microbiol.* 9:4091-4095.
- Costa, P. I., Franco, C. F., Miranda, V. S., Teixeira, D. C., and Hartung, J. S. 2000. Strains of *Xylella fastidiosa* rapidly distinguished by arbitrarily primed-PCR. *Curr. Microbiol.* 40:279-282.
- Dietrich, W. F., Weber, J. L., Nickerson, D. A., and Kwok, P. 1998. Identification and analysis of DNA polymorphisms. Pages 135-186 in: *Genome Analysis: A Laboratory Manual*. CSHL Press, Woodbury, NY.
- Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783-791.
- Fry, S. M., Huang, J. S., and Milholland, R. D. 1994. Isolation and preliminary characterization of extracellular proteases produced by strains of *Xylella fastidiosa* from grapevines. *Phytopathology* 84:357-363.
- Gordon, D., Abajian, C., and Green, P. 1998. Consed: a graphical tool for sequence finishing. *Genome Res.* 8:195-202.
- Gordon, D. M., Wexler, M., Reardon, T. B., and Murphy, P. J. 1995. The genetic structure of *Rhizobium* populations. *Soil Biol. Biochem.* 27:491-499.
- Hartmann, J. L., IV, Garvik, B., and Hartwell, L. 2001. Principles for the buffering of genetic variation. *Science* 291:1001-1004.
- 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.
- Hopkins, D. L. 1984. Variability of virulence in grapevine among isolates of Pierce's disease bacterium. *Phytopathology* 74:1395-1398.
- Hopkins, D. L. 1989. *Xylella fastidiosa*: xylem-limited bacterial pathogen of plants. *Annu. Rev. Phytopathol.* 27:271-290.
- Hopkins, D. L., and Adlerz, W. C. 1988. Natural hosts of *Xylella fastidiosa* in Florida. *Plant Dis.* 72:429-431.
- Hopkins, D. L., Thompson, C. M., and Blistine, F. E. 1989. Relationships between xylem-limited bacteria and citrus blight. *Proc. Fla. State Hort. Soc.* 102:21-23.
- Huang, P.-Y., Milholland, R. D., and Daykin, M. E. 1986. Structural and morphological changes associated with the Pierce's disease bacterium in bunch and muscadine grape tissue. *Phytopathology* 76:1232-1238.
- Kamper, S. M., French, W. J., and Dekloet, S. R. 1985. Genetic relationship of some fastidious xylem-limited bacteria. *Int. J. Syst. Bacteriol.* 35:185-188.
- Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16:111-120.
- Kumar, S., Tamura, K., and Nei, M. 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief. Bioinform.* 5:150-163.
- Lemos, E. G. M., Carareto-Alves, L. M., and Campanharo, J. C. 2003. Genomics-based design of defined growth media for the plant pathogen *Xylella fastidiosa*. *FEMS Microbiol. Lett.* 219:39-45.
- Leu, L. S., and Su, C. C. 1993. Isolation, cultivation and pathogenicity of *X. fastidiosa*, the causal bacterium of pear leaf scorch disease in Taiwan. *Plant Dis.* 77:642-646.
- Leung, H., Nelson, R. J., and Leach, J. E. 1993. Population structure of plant pathogenic fungi and bacteria. *Adv. Plant Pathol.* 10:157-250.
- Lima, J. E. O., Miranda, V. S., Hartung, J. S., Bransky, R. H., Coutinho, A., Roberto, S. R., and Carlos, E. F. 1998. Coffee leaf scorch bacterium: Axenic culture, pathogenicity, and comparison with *Xylella fastidiosa* of citrus. *Plant Dis.* 82:94-97.
- Martin, A. P. 2002. Phylogenetic approaches for describing and comparing the diversity of microbial communities. *Appl. Environ. Microbiol.* 68:3673-3682.
- Maynard Smith, J., Smith, N. H., O'Rourke, M., and Spratt, B. G. 1993. How clonal are bacteria? *PNAS* 90:4384-4388.
- Mehta, A., Leite, R. P., and Rosato, Y. B. 2001. Assessment of the genetic diversity of *Xylella fastidiosa* isolated from citrus in Brazil by PCR-RFLP of 16S-23S intergenic spacer and rep-PCR fingerprinting. *Antonie Leeuwenhoek* 79:53-59.
- Minsavage, G. V. 1994. Development of a polymerase chain reaction protocol for detection of *Xylella fastidiosa* in plant tissue. *Phytopathology* 84:456-461.

33. Morin, P. A., Luikart, G., Wayne, R. K., and the SNP workshop group. 2004. SNPs in ecology, evolution and conservation. *Trends Ecol. Evol.* 19:208-216.
34. Nei, M. 1972. Genetic distance between populations. *Am. Nat.* 106:283-292.
35. Nei, M., and Kumar, S. 2000. *Molecular Evolution and Phylogenetics*. Oxford University Press, Oxford, UK.
36. Paradela Filho, O., Sugimori, M. H., Ribeiro, I. J. A., Garcia, A., Beretta, M. J. G., Harakawa, R., Machado, M. A., Laranjeira, F., Rodrigues Neto, J., and Beriam, L. O. S. 1997. Constatação de *Xylella fastidiosa* em cafeeiro no Brasil. *Summa Phytopatol.* 23:46-49.
37. Pooler M. R., and Hartung, J. S. 1995. Genetic relationships among strains of *Xylella fastidiosa* from RAPD-PCR data. *Curr. Microbiol.* 31:134-137.
38. Pooler M. R., and Hartung, J. S. 1995. Specific PCR detection and identification of *Xylella fastidiosa* strains causing citrus variegated chlorosis. *Curr. Microbiol.* 31:377-381.
39. Qin, X., Miranda, V. S., Machado, M. A., Lemos, E. G. M., and Hartung, J. S. 2001. An evaluation of the genetic diversity of *Xylella fastidiosa* isolated from citrus and coffee in São Paulo, Brazil. *Phytopathology* 91:599-605.
40. Rosato, Y. B., Neto, J. R., Miranda, V. S., Carlos, E. F., and Manfio, G. P. 1998. Diversity of a *Xylella fastidiosa* isolated from *Citrus sinensis* affected by citrus variegated chlorosis in Brazil. *Syst. App. Microbiol.* 21:593-598.
41. Saitou, N., and Nei, M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4:406-425.
42. Sambrook, J., Maniatis, T., and Fritsh, E. F. 1989. *Molecular Cloning: A Laboratory Manual*, 2nd ed. Cold Sprig Harbor Laboratory, Cold Spring Harbor, NY.
43. Sanderlin, R. S. 1998. Evidence that *Xylella fastidiosa* is associated with pecan fungal leaf scorch. *Plant Dis.* 82:264-270.
44. Schaad, N. W., Postnikova, E., Lacy, G., Fatmi, M. B., 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.
45. Schneider, S., Roeslli, D., and Excoffier, L. 2000. ARLEQUIN: A software for population genetics data analysis. Version 2.000. University of Geneva, Switzerland.
46. 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.
47. Simpson, A. J. G., Reinach, F. C., and Arruda, P. 2000. The genome sequence of plant pathogen *Xylella fastidiosa*. *Nature* 406:151-159.
48. Spiers, A. J., Buckling, A., and Rainey, P. B. 2000. The causes of *Pseudomonas* diversity. *Microbiology* 146:2345-2350.
49. Stoneking, M. 2001. Single nucleotide polymorphism: From the evolutionary past.... *Nature* 409:821-822.
50. Wells, J. M., Raju, B. C., Jung, H. Y., Weinsburg, W. G., Mandelco-Paul, L., and Brenner, D. J. 1987. *Xylella fastidiosa*: gen. nov., sp. nov., gram-negative, xylem-limited fastidious plant bacteria related to *Xanthomonas* spp. *Int. J. Syst. Bacteriol.* 37:136-143.
51. Zhu, Y., Chen, H., Fan, J., Wang, Y., Lee, Y., Chen, J., Fan, J., Yang, S., Hee, L., Zeung, H., Mew, T. W., and Mundt, C. C. 2000. Genetic diversity and disease control in rice. *Nature* 406:718-722.