Journal of Bacteriology

Comparative Analyses of the Complete Genome Sequences of Pierce's Disease and Citrus Variegated Chlorosis Strains of *Xylella fastidiosa*

M. A. Van Sluys, M. C. de Oliveira, C. B. Monteiro-Vitorello, C. Y. Miyaki, L. R. Furlan, L. E. A. Camargo, A. C. R. da Silva, D. H. Moon, M. A. Takita, E. G. M. Lemos, M. A. Machado, M. I. T. Ferro, F. R. da Silva, M. H. S. Goldman, G. H. Goldman, M. V. F. Lemos, H. El-Dorry, S. M. Tsai, H. Carrer, D. M. Carraro, R. C. de Oliveira, L. R. Nunes, W. J. Siqueira, L. L. Coutinho, E. T. Kimura, E. S. Ferro, R. Harakava, E. E. Kuramae, C. L. Marino, E. Giglioti, I. L. Abreu, L. M. C. Alves, A. M. do Amaral, G. S. Baia, S. R. Blanco, M. S. Brito, F. S. Cannavan, A. V. Celestino, A. F. da Cunha, R. C. Fenille, J. A. Ferro, E. F. Formighieri, L. T. Kishi, S. G. Leoni, A. R. Oliveira, V. E. Rosa Jr., F. T. Sassaki, J. A. D. Sena, A. A. de Souza, D. Truffi, F. Tsukumo, G. M. Yanai, L. G. Zaros, E. L. Civerolo, A. J. G. Simpson, N. F. Almeida Jr., J. C. Setubal and J. P. Kitajima

Updated information and services can be found at:
http://jb.asm.org/content/185/3/1018

These include:

REFERENCES

This article cites 37 articles, 13 of which can be accessed free at: http://jb.asm.org/content/185/3/1018#ref-list-1

CONTENT ALERTS

Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), more»

Information about commercial reprint orders: **http://journals.asm.org/site/misc/reprints.xhtml**
To subscribe to to another ASM Journal go to: **http://journals.asm.org/site/subscriptions/**

Journals.ASM.org

Comparative Analyses of the Complete Genome Sequences of Pierce's Disease and Citrus Variegated Chlorosis Strains of *Xylella fastidiosa*

M. A. Van Sluys,[1]* M. C. de Oliveira,[1] C. B. Monteiro-Vitorello,[2] C. Y. Miyaki,[1] L. R. Furlan,[3]†
L. E. A. Camargo,[2] A. C. R. da Silva,[4]† D. H. Moon,[5] M. A. Takita,[6] E. G. M. Lemos,[7]
M. A. Machado,[6] M. I. T. Ferro,[7] F. R. da Silva,[8] M. H. S. Goldman,[9] G. H. Goldman,[10]
M. V. F. Lemos,[7] H. El-Dorry,[4] S. M. Tsai,[5] H. Carrer,[2] D. M. Carraro,[11] R. C. de Oliveira,[12]
L. R. Nunes,[12] W. J. Siqueira,[13] L. L. Coutinho,[2] E. T. Kimura,[14] E. S. Ferro,[14] R. Harakava,[15]
E. E. Kuramae,[16] C. L. Marino,[17] E. Giglioti,[18] I. L. Abreu,[7] L. M. C. Alves,[7] A. M. do Amaral,[6]†
G. S. Baia,[14] S. R. Blanco,[4] M. S. Brito,[9] F. S. Cannavan,[5] A. V. Celestino,[13] A. F. da Cunha,[19]
R. C. Fenille,[16] J. A. Ferro,[7]† E. F. Formighieri,[5] L. T. Kishi,[7] S. G. Leoni,[14] A. R. Oliveira,[1]
V. E. Rosa Jr.,[8] F. T. Sassaki,[17] J. A. D. Sena,[7] A. A. de Souza,[6]† D. Truffi,[2] F. Tsukumo,[19]
G. M. Yanai,[12] L. G. Zaros,[2] E. L. Civerolo,[20] A. J. G. Simpson,[11]
N. F. Almeida Jr.,[21] J. C. Setubal,[22] and J. P. Kitajima[8,22]‡

Instituto de Biociências,[1] Instituto de Química,[4] and Instituto de Ciências Biomédicas,[14] Universidade de São Paulo, 05508-900 São Paulo, Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, 13418-900, Piracicaba,[2] Centro de Energia Nuclear na Agricultura, Universidade de São Paulo, 13400-970, Piracicaba,[5] Centro de Citricultura Sylvio Moreira, Instituto Agronômico de Campinas, 13490-970, Cordeirópolis,[6] Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, 14884-900, Jaboticabal,[7] Centro de Biologia Molecular e Engenharia Genética,[8] Instituto de Biologia,[19] and Instituto de Computação,[22] Universidade Estadual de Campinas, 13083-970, Campinas, Faculdade de Filosofia, Ciências, e Letras[9] and Faculdade de Ciencias Farmaceuticas,[10] Universidade de São Paulo, 14040-901, Ribeirão Preto, Instituto Ludwig de Pesquisa Sobre o Câncer, 01509-010, São Paulo,[11] Núcleo Integrado de Biotecnologia, Universidade de Mogi das Cruzes, 08780-911, Mogi das Cruzes,[12] Instituto Agronômico de Campinas, 13001-970, Campinas,[13] Instituto Biológico, 04014-002, São Paulo,[15] Faculdade de Medicina Veterinária e Zootecnia[3] and Faculdade de Ciências Agronômicas,[16] Universidade Estadual Paulista, 18603-970, Botucatu, Instituto de Biociências, Universidade Estadual Paulista, 18618-000, Botucatu,[17] Universidade Federal de São Carlos, 13600-970, Araras, São Paulo,[18] and Departamento de Computação e Estatistica, Universidade Federal de Mato Grosso do Sul, 79070-900, Campo Grande, Mato Grosso do Sul,[21] Brazil, and Agricultural Research Service, U.S. Department of Agriculture, Parlier, California 93648[20]

Xylella fastidiosa is a xylem-dwelling, insect-transmitted, gamma-proteobacterium that causes diseases in many plants, including grapevine, citrus, periwinkle, almond, oleander, and coffee. *X. fastidiosa* has an unusually broad host range, has an extensive geographical distribution throughout the American continent, and induces diverse disease phenotypes. Previous molecular analyses indicated three distinct groups of *X. fastidiosa* isolates that were expected to be genetically divergent. Here we report the genome sequence of *X. fastidiosa* (Temecula strain), isolated from a naturally infected grapevine with Pierce's disease (PD) in a wine-grape-growing region of California. Comparative analyses with a previously sequenced *X. fastidiosa* strain responsible for citrus variegated chlorosis (CVC) revealed that 98% of the PD *X. fastidiosa* Temecula genes are shared with the CVC *X. fastidiosa* strain 9a5c genes. Furthermore, the average amino acid identity of the open reading frames in the strains is 95.7%. Genomic differences are limited to phage-associated chromosomal rearrangements and deletions that also account for the strain-specific genes present in each genome. Genomic islands, one in each genome, were identified, and their presence in other *X. fastidiosa* strains was analyzed. We conclude that these two organisms have identical metabolic functions and are likely to use a common set of genes in plant colonization and pathogenesis, permitting convergence of functional genomic strategies.

Different microorganisms are able to survive in and to colonize plant water-conductive vessels (xylem). The result of this association is either beneficial or detrimental to the plant host.

Of the latter, an example is the association of *Xylella fastidiosa* (38) with diverse plant hosts. *X. fastidiosa* is a fastidious, insect-transmitted, xylem-inhabiting bacterium known to cause several economically important diseases of both monocotyledonous and dicotyledonous plants (14, 17, 29). These diseases include Pierce's disease (PD) of grapevine and citrus variegated chlorosis (CVC), which have rather distinct symptoms and geographical distributions.

PD, caused by certain strains of *X. fastidiosa*, is characterized by wilted, shriveled, raisin-like fruit and scorched leaves that detach, leaving bare petioles attached to the canes (37). The bark of affected canes may lignify or mature irregularly, leaving

---

* Corresponding author. Mailing address: Departamento de Botânica, Instituto de Biociências, Universidade de São Paulo, R. do Matão, 277, 05508-900, São Paulo, São Paulo, Brazil. Phone: 55(11)30917548. Fax: 55(11)30917724. E-mail: mavsluys@usp.br.
† Present address: Empresa Brasileira de Pesquisa Agropecuaria, Cordeiropolis, Brazil.
‡ Present address: Alellyx Applied Genomics, 13067-850, Campinas, São Paulo, Brazil.

TABLE 1. General genome features of *X. fastidiosa* Temecula, causal agent of PD, and *X. fastidiosa* 9a5c, causal agent of CVC[a]

Disease	Length (bp)	GC content (%)	Coding region (% chromosome size)	No. of predicted ORFs	Avg ORF length (bp)	% of:			No. of:			Plasmid length (bp)
						ORFs with a functional assignment	Conserved hypothetical ORFs	Hypothetical ORFs	tRNAs	rRNA operons	tmRNAs	
PD[b]	2,519,802	51.8	79.9	2,066	979	65.9	33.3	0.8	49	2	1	1,345
CVC[c]	2,679,305	52.7	80.6	2,249	964	57.4	38.2	4.4	49	2	1	1,285 and 51,158

[a] Temecula was isolated from grapevine in California, and 9a5c was isolated from citrus in São Paulo, Brazil (33).
[b] See supplementary material mentioned in the text for gene maps, gene lists, and categories.
[c] The data presented are from our updated genome analysis. This revision removed small hypothetical ORFs, now thought to be false predictions.

areas of brown bark tissue surrounded by green immature tissue. Delayed and stunted shoot growth occurs in the spring following infection, and chronically infected grapevines eventually die. This devastating disease is a major threat to the viability of the California wine industry. The PD *X. fastidiosa* strain, whose genome sequence is described here, was isolated in 1998 from a naturally infected grapevine in Temecula, Calif. CVC, on the other hand, is characterized by the presence of small hard fruits of no commercial value and conspicuous spotted chlorosis on the upper leaf surface, resembling the symptoms of zinc deficiency and occasionally accompanied by gum-like extrusions from the spots on the lower surface (4).

X. fastidiosa is phylogenetically placed at the base of the gamma group of *Proteobacteria* (36). Molecular analyses at the species level have revealed three distinct groups. The grapevine-infecting variants responsible for PD are found in one group, while the citrus-infecting variants responsible for CVC are found in another (6, 7, 16). Initial expectations, based on geographical distributions, host diversity, differential disease symptoms, and molecular analyses, were that organisms from the three groups would be sufficiently different to support taxonomic separation at the subspecies or pathovar level (17, 19, 24).

The genomic sequence of the PD *X. fastidiosa* Temecula strain has now been determined in order to further elucidate both the molecular basis of *X. fastidiosa* pathogenicity and the phylogenetic relationships among *X. fastidiosa* strains. Comparative analyses of the complete sequences and annotations of PD *X. fastidiosa* and an *X. fastidiosa* representative of the CVC group isolated from Brazil (33) revealed that these strains exhibit remarkably limited genomic variability and share 95.7% amino acid identity in equivalent regions. There are only three genomic rearrangements, two identified genomic islands, 41 PD *X. fastidiosa* and 152 CVC *X. fastidiosa* strain-specific genes, and some genes harboring frameshifts. Our analyses suggest that a common functional genomic strategy may be undertaken to identify means of controlling *X. fastidiosa*-induced diseases.

MATERIALS AND METHODS

Genome sequence and assembly. Total genomic DNA was isolated from the PD *X. fastidiosa* Temecula strain. The complete genome sequence was generated by using a combination of ordered cosmid and shotgun strategies (13). Various shotgun libraries with different insert sizes (0.8 to 2.0 kb and 2.0 to 4.5 kb) were constructed from nebulized genomic DNA cloned into pUC18, and a total of 102,348 sequences were generated; 81% of these had at least 400 bases with a Phred quality above 20 (15), providing approximately 13-fold genome coverage. A cosmid library (Lawrist vector) with inserts ranging from 30 to 45 kb was constructed. A total of 2,752 cosmid ends were sequenced; 63% of these had at

least 300 bases with a Phred quality above 20, providing approximately 26-fold genome coverage. These cosmid ends were used in the scaffold, and 12 cosmids were selected to be fully sequenced. Sequence gaps were identified by linking information from forward and reverse reads and were closed by primer walking, PCR sequencing, and insert subcloning. Sequences from both ends of most cosmid clones were used to confirm the orientation and integrity of the contigs. The sequences were assembled by using the Phred+Phrap+Consed package (15). All consensus bases have a Phred quality of at least 20. There are no unexplained high-quality discrepancies, and the overall error estimate is less than 1 in every 10,000 bases. Most of the sequencing was performed with BigDye terminators and ABI Prism 3700 DNA sequencers.

Genome annotation and analysis. Annotation was dependent primarily on open reading frame (ORF) identification by using GLIMMER (10), GeneMark (5), and alignment against the National Center for Biotechnology Information protein database. BLASTX searches were carried out to find additional putative protein-coding genes. All ORFs were inspected manually by the annotation team. For each ORF, links to Cluster of Orthologous Groups of Proteins (COG), Protein Family Database (PFAM), and Kyoto Encyclopedia of Genes and Genomes (KEGG) were made available. RNA species were identified by using BLASTN (3), secondary structure analysis, and tRNAscan-SE (21). Domestic software was used in order to generate items such as gene maps, lists, comparative CVC *X. fastidiosa* and PD *X. fastidiosa* data, and GenBank submissions. For a full list of ORFs, gene maps, and comparative tables, refer to supplementary material at http://aeg.lbi.ic.unicamp.br/world/xfpd/.

Whole genomes of PD *X. fastidiosa* and CVC *X. fastidiosa* were compared at the nucleotide level by using the program MUMmer (11) with default values. At the amino acid level, the genomes were compared by using previously developed programs (33). Genes *g* and *h* were considered orthologs if *h* was the best BLASTP hit for *g* and vice versa, where the e-values were 10^{-5} or less. A gene was considered strain specific if it had no hits or the e-value was 10^{-5} or more in the other genome.

PCR analysis of genomic islands in different strains of *X. fastidiosa*. Oligonucleotides were constructed for the genomic islands and their flanking regions. PCR analyses were carried out in duplicate in two different laboratories. For a full list of strains and primers, refer to supplementary material at http://aeg.l-bi.ic.unicamp.br/world/xfpd/.

Nucleotide sequence accession numbers. The sequences have been deposited in GenBank with accession numbers AE009442 (chromosome) and AE009443 (plasmid).

RESULTS AND DISCUSSION

Genome features. The PD *X. fastidiosa* Temecula genome is composed of a single large circular chromosome (2,519,802 bp) and a small plasmid, pXFPD1.3 (1,345 bp), also reported by others for some PD *X. fastidiosa* strains (16). Table 1 shows a comparative summary of the main genome features of PD *X. fastidiosa* Temecula and CVC *X. fastidiosa* 9a5c. Major discrepancies between these two strains consist of a 159,503-bp chromosome size difference and the absence of large plasmid pXF51 in PD *X. fastidiosa* Temecula. The variation in the percentages of hypothetical ORFs observed (0.8% for PD *X. fastidiosa* Temecula and 4.4% for CVC *X. fastidiosa* 9a5c) could be due to the difference in the genome size, as explained below.
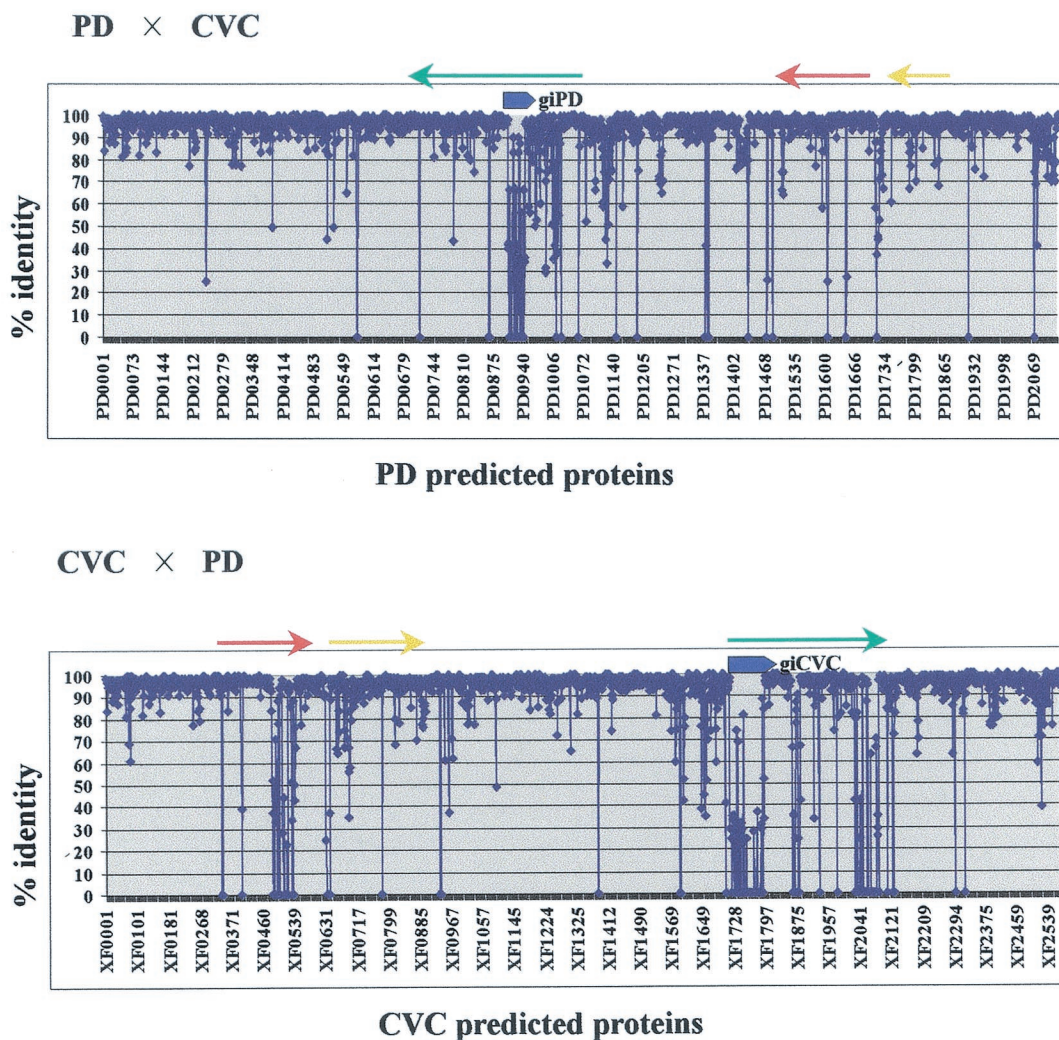
PD × CVC

CVC × PD

FIG. 1. Global comparison of PD *X. fastidiosa* Temcula and CVC *X. fastidiosa* 9a5c proteomes. The distribution of protein-coding gene sequence identity along each genome is shown. Amino acid identity was determined by a reciprocal BLAST analysis of one proteome against the other. The positions of the genomic islands (giPD1 and giCVC1) are illustrated at the top of each genome graph. The colored arrows at the top of each graph indicate the positions of rearranged sequences (see Fig. 2B).

Of the 2,066 protein-coding genes annotated in PD *X. fastidiosa* Temecula, 2,025 (98%) are also present in CVC *X. fastidiosa* 9a5c. Of these orthologous genes, 94.5% have 80% or more amino acid identity, with an average identity of 95.7%, as shown in Fig. 1. This conservation is distributed along the whole chromosome, and regions of lower identity tend to appear in clusters (Fig. 1). This level of protein identity is comparable to that observed among the orthologous proteins of different *Escherichia coli* strains (27), *Helicobacter pylori* strains (1), and *Salmonella enterica* serovars (22) and thus supports a close relationship between these two *Xylella* strains. The most conserved PD *X. fastidiosa* Temecula genes include all those that determine the basic metabolism and cellular functions of the bacterium, which we thus conclude are mostly identical to those previously described for CVC *X. fastidiosa* 9a5c (33). Energy is generated by the efficient utilization of carbohydrates, including cellulose, but with no predicted catabolism of fatty acids or amino acids as alternative energy sources. In contrast, a complete set of biosynthetic pathways is present,

permitting the synthesis of all amino acids, purines, pyrimidines, and nucleotides as well as an extensive array of cofactors and prosthetic groups. Transport systems include those for carbohydrates, ions, amino acids, and peptides as well as those for the extrusion of drugs and toxins.

A total of 106 genes in the PD *X. fastidiosa* Temecula genome (5.2%), although shared with the CVC *X. fastidiosa* 9a5c genome, have amino acid identities of 20 to 80%. Among these are 58 genes that are found within phage-related regions and genomic islands. In addition, 18 conserved hypothetical genes in the vicinity of the hemolysin and hemagglutinin genes fall within this group. Interestingly, among genes with assigned functions that exhibited this higher level of divergence, we found some that may be involved in *X. fastidiosa*-plant host interactions, including those for fimbrillins and hemagglutinins (attachment and cell aggregation); colicin, hemolysin, and bacteriocin (toxins); and drug resistance and DNA restriction and modification enzymes (see supplementary material for a full list of genes). Thus, there may have been more selective pres-

TABLE 2. Strain-specific genes and genes with a frameshift and/or a stop codon in frame

Strain	No. and type of:	
	Strain-specific genes	Genes with frameshift and/or stop codon in frame
Temecula	16 Hypothetical 16 Conserved hypothetical 13 Mobile genetic elements 2 DNA methyltransferases 1 Type II restriction enzyme 1 HicA-related protein 1 Hydrolase 1 Integral membrane protein	7 Conserved hypothetical 10 Mobile genetic elements 2 Hemagglutinin-secreted proteins 1 Fimbrial adhesin 3 Type I restriction-modification enzymes (R, S, and M) 1 DNA methyltransferase 1 Aconitate hydratase 1 Mg^{2+} transporter 1 Heat shock protein (HslU) 1 Anthranilate synthase component I 1 Penicillin binding protein
9a5c	99 Hypothetical 21 Conserved hypothetical 19 Mobile genetic elements 2 DNA methyltransferases 1 Single-stranded DNA binding protein 1 O antigen acetylase 1 Transport protein 1 NADH flavin oxidoreductase 1 Oxidoreductase 1 HTP reductase 4 Transcriptional regulators 1 Drug resistance translocase	2 Conserved hypothetical 7 Mobile genetic elements 1 Integral membrane protein 1 Polygalacturonase precursor 2 Type I restriction-modification enzymes (R and M) 1 DNA damage-inducible protein 1 Aconitate hydratase 1 Ribosomal protein S6 modification protein

sure for alterations in these genes to enhance plant-specific bacterial colonization capability. There are also genes in the two genomes that have either a frameshift or an in-frame stop codon (Table 2), suggesting that they are nonfunctional. The most intriguing of these is the polygalacturonase precursor gene, which has a stop codon in CVC *X. fastidiosa* 9a5c but is intact in PD *X. fastidiosa* Temecula. For two other partially sequenced *Xylella* genomes (http://www.jgi.doe.gov/), no frameshift is observed within the polygalacturonase precursor gene. Other than among *Xylella* genomes, this gene shares 65% identity with its *Ralstonia solanacearum* ortholog (32); orthologs are also present in other necrogenic plant pathogens, such as *Xanthomonas campestris* pv. *campestris*, *X. axonopodis* pv. *citri*, and *Erwinia carotova*. This gene is essential for the synthesis of cell wall-degrading enzymes that facilitate intervessel migration. Its intact status in PD *X. fastidiosa* Temecula may account for the more aggressive nature of PD than of CVC (2), where it is not essential for disease development, since Koch's postulates for strain 9a5c were experimentally fulfilled.

PD *X. fastidiosa* Temecula has 41 strain-specific genes (1.9%), while CVC *X. fastidiosa* 9a5c has 152 such genes (6.8%) (Table 2). In both strains, more than half of these are hypothetical or conserved hypothetical genes, and a significant proportion are associated with mobile genetic elements. Among the PD *X. fastidiosa*-specific genes with assigned functions are a hydrolase gene with similarity to genes in *Xanthomonas* (gi21113352), *Pseudomonas* (gi15598992), and *Salmonella* (gi16763691) and a gene for a type II restriction and modification system most similar to that of the cyanobacterium *Nostoc* (gi547934). Genes for two other proteins, proteic killer and HicA, are shared with *Nostoc* (gi17232769) and *E. coli* (gi15804020). The CVC *X. fastidiosa*-specific genes with assigned functions include a gene for an O-antigen acetylase that is involved in LPS modification and that is similar to

those found in *S. enterica* (gi16761319), *Sinorhizobium meliloti* (gi16761319), *Mesorhizobium loti* (gi13474718), *Neisseria meningitidis* (gi15795071), and *Pseudomonas* (gi15600431) (12, 34) and an additional drug resistance translocase gene that is most similar to genes identified in *Caulobacter crescentus* (gi16127299), *Mycobacterium tuberculosis* (gi15841836), and *M. loti* (gi13472297). This gene is located, along with 71 other specific genes, on the CVC-specific island described below. The apparently diverse origins of the specific genes with assigned functions in the two strains are also reflected in conserved hypothetical genes that are similar to the genes of a large, unrelated group of bacteria including *Xanthomonas*, *Pseudomonas*, *Ralstonia*, *Listeria innocua*, *Agrobacterium tumefaciens* C58, *C. crescentus*, and even, for one gene in CVC *X. fastidiosa*, the distantly related eubacterium *Microscilla* (gi14485002). It appears that individual genes have been accumulated in phages and transferable islands during their passage through many bacterial species before being incorporated within the *X. fastidiosa* genome (26).

Genome reorganization. Alignment of the PD *X. fastidiosa* Temecula and CVC *X. fastidiosa* 9a5c chromosomes, starting from the putative origins of replication, highlighted three chromosomal regions of the two genomes that were translocated and inverted despite their overall identity (Fig. 2). All such reorganization events occurred at least 250,000 bp from the putative origin of replication (Fig. 2A), as previously observed for other bacterial species (18). These three large rearranged chromosomal regions and other small rearrangements were all flanked at one border by a putative phage-related integrase, suggesting that they were phage mediated. The PD *X. fastidiosa* Temecula chromosome harbors eight clusters of phage-related regions, Xpd1 to Xpd8, none of which is organized in a manner similar to that of the four CVC *X. fastidiosa* 9a5c prophages described previously (XfP1 to XfP4) (33). The
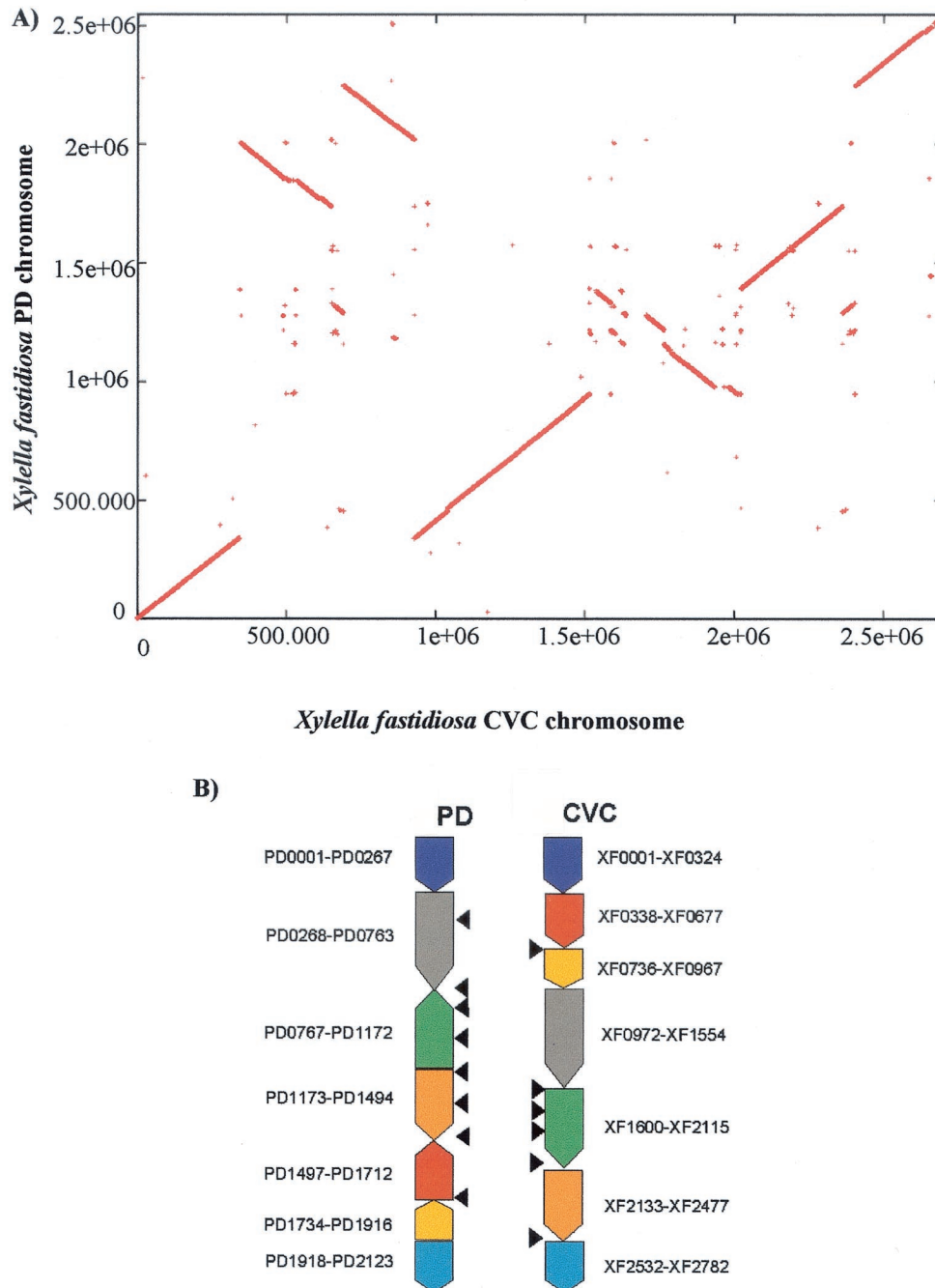
FIG. 2. Chromosome alignment of PD *X. fastidiosa* Temecula and CVC *X. fastidiosa* 9a5c. (A) Nucleotide sequence alignment of both genomes, starting at the putative origins of replication, as determined with MUMmer. (B) Chromosome backbones of both genomes, showing the relative positions, sizes, and orientations of colinear chromosome regions detected in panel A. The direction of the arrow within each chromosome fragment indicates its relative orientation. Black triangles illustrate the positions of phage-related regions and genomic islands within each chromosome.

Stretcher global alignment program (25) was used to determine the overall nucleotide identity in a given region, enabling analysis of similarity among the phage-related regions. The Xpd1 region shares 83 and 78% nucleotide identity with CVC *X. fastidiosa* 9a5c prophages XfP2 and XfP1, respectively. All of the other CVC *X. fastidiosa* 9a5c prophage regions share less than 50% nucleotide identity with the Xpd phage clusters. Three of the phage-related regions are specific to the PD *X.*

fastidiosa genome, and one is involved in one of the large rearrangements mentioned above. Three phage-related regions (Xpd5, Xpd6, and Xpd8) are highly divergent from the equivalent regions in the CVC *X. fastidiosa* chromosome, while Xpd1 maintains the same borders as XfP4. We have not considered these phage-related regions to be strain specific due to the fact that we cannot determine whether the insertion events occurred prior to strain divergence. In addition, some genes

are shared by these phage-related regions. Figure 2B is a schematic representation of the PD *X. fastidiosa* and CVC *X. fastidiosa* chromosomes illustrating the rearrangements and the relative positions of the prophage clusters.

Genomic islands. Genomic islands specific to each genome were characterized on the basis of marked decreases in protein identities, different GC contents, and codon bias. Two of these islands, one specific to each genome, have higher GC contents, and their relative positions are indicated in Fig. 1. In PD *X. fastidiosa* Temecula, genomic island PD1 (giPD1) is 15.7 kb long, has 61.2% GC content, and harbors an extra copy of a hemagglutinin gene with a phage-related integrase at one end (Fig. 3A). In CVC *X. fastidiosa* 9a5c, genomic island CVC1 (giCVC1) is 67 kb long, has 63.3% GC content, and is inserted within tRNA Gly-2 (Fig. 3B). The integrase immediately adjacent to the tRNA Gly-2 gene is highly similar at the nucleotide (93%) and protein (87%) levels to a previously described *Pseudomonas putida* strain B13 integrase (31) that is associated with the P2 integrase/recombinase family. In *Pseudomonas*, the integrase is associated with a self-transmissible 105-kb *clc* element that carries the *clcRABDE* genes encoding chlorocatechol-degradative enzymes. It is interesting that different integrases can share common integration target sites (39). The integrase characterized for giCVC1 is targeted to the glycine tRNA structural gene (*glyV*), like the integrase associated with the *Pseudomonas* self-transmissible element.

In an attempt to correlate the presence or absence of the genomic islands with a disease phenotype, PCR analyses were performed to characterize their distributions in different *X. fastidiosa* strains. Primers were constructed for the island borders. Table 3 shows the giCVC1 distribution in 64 strains of *X. fastidiosa* isolated from different hosts and different geographical regions. The flanking regions of giCVC1 are the same as the corresponding regions in the PD *X. fastidiosa* Temecula genome, except that the whole region is inverted relative to the origin of replication. The use of the same flanking primers produced, after PCR amplification, a product of 4,587 bp for the PD *X. fastidiosa* genome, indicating the absence of giCVC1. However, three distinct groups were identified based on the sizes of the amplified products (Table 3). The CVC group, which contains all or part of giCVC1, comprises most of the tested Brazilian strains regardless of host (citrus, coffee, hibiscus, and periwinkle). Surprisingly, the PD group, which does not have giCVC1, was subdivided into two groups. One group comprises strains isolated from grapevine, mulberry, almond (ATCC 35870), and oleander, with an amplified fragment of approximately 4.6 kb. The second group, comprising strains isolated from plum in Brazil and the United States and from almond (ATCC 700965), elm, oak, and periwinkle in the United States, produced a smaller amplified fragment (2.9 kb). These results are consistent with the existence of different groups of strains of *Xylella* in North America and South America, as suggested previously (8, 28), with the exception of the plum strain isolated in Brazil. This strain had a pattern similar to that of the North American strains, a fact that could be indicative of its recent introduction into Brazil via infected seedlings.

In PD *X. fastidiosa*, giPD1 is located within the phage-related region Xpd2 (Fig. 3A). PCR analysis of this region in 30 different *X. fastidiosa* strains revealed a pattern more variable than that obtained for the giCVC1 distribution. The presence or absence of giPD1 could not be correlated with the groups described above, as both PD and CVC strains may contain this island. Careful inspection of the genome around giPD1 enabled us to characterize a 68.8-kb region that could represent the ancient insertion of a prophage and/or a conjugative transposon related to the Tn*21* family. Transposon ends similar to those of Tn*5053* were detected at both extremities of the proposed region, and a degenerate copy of the transposase was also found within the island. Tn*5053* was originally described as a transposon which carries the mercuric resistance operon described for *Xanthomonas* (20). One interpretation of the PCR results is that giPD1 was already present in the ancestral *Xylella* genome prior to the divergence of the 30 strains studied here and that the evolution of each strain, irrespective of the plant host, was characterized by multiple losses from this ancestral island. On the other hand, it seems that giCVC1 is limited to Brazilian *X. fastidiosa* strains. Therefore, it is reasonable to infer that this island is a recent acquisition by the original strain that spread to South America that nevertheless occurred prior to its expansion. It is interesting that part of giCVC1 was also observed in *X. axonopodis* pv. *citri* (9), which causes citrus canker.

The biology of *Xylella*-induced disease is poorly understood, and the grouping of strains has been a strategy devised in part to develop effective disease management measures. Different techniques have been used to try to establish pathovar or subspecies categories not necessarily focused on the evolution of the group. Qin et al. (30) proposed the following natural groups of strains: 1, citrus and coffee; 2, grapevine, almond, and ragweed; and 3, elm, oak, and plum. Chen et al. (6) proposed three strain groups based on 16S ribosomal DNA sequences: 1, citrus and coffee; 2, grapevine and mulberry; and 3, elm, oak, peach, plum, and periwinkle. Based on our analysis of giCVC1, we propose three groups of strains: 1, citrus, coffee, and possibly other South American strains; 2, grapevine, mulberry, oleander, and some almond strains, all from North America; and 3, elm, oak, plum, periwinkle, and some almond strains, again all from North America. Hendson et al. (16) showed that with the exception of some almond strains, all *X. fastidiosa* strains isolated from the same host had identical sequences for the intergenic spacer of rRNA genes. The two almond strains that we examined are distinguished by the size of the amplified fragment corresponding to giCVC1 and by the presence of giPD1.

The presence of genomic islands in different closely related strains is known to represent the gain of adaptive traits by an organism; examples include the *Mesorhizobium* symbiotic island (35) and the LEE islands of enteropathogenic *E. coli* strains (23). The acquisition of such islands can result in evolution by quantum leaps. A comparative analysis of the two *X. fastidiosa* genomes permitted the identification of such islands in both genomes, although their adaptive functions remain to be demonstrated. Essentially all of the differences between the PD *X. fastidiosa* Temecula and CVC *X. fastidiosa* 9a5c genomes can be accounted for by the numbers and relative positions of clusters of phage-related genes and insertion or deletion events, among which giPD1 and giCVC1 are included. If

A) Left and right junctions of the giPD in   PD Temecula and corresponding region in   CVC 9a5c

```
CVC:  1607753  TTGGAGTCGCTGGGTGTGAGTGTCGCACGCGCTCCAGACCGGGCGGCGCCACGCGCGGGCGGCCGCTGCCCGACTGGGCGAT  1607873
PD:   1196420  TTGGAGTCGCTGGGTGTGAGTGTCGCACGCGCTCCAGACCGGGCGGCGCCACGCGCGGGCGGCCGCTGCCCGGATTGGGCGAT  1196300

CVC:  1607874  CAGGCGCGACAAGATGAGCTATGGGCGGTACCTGCTGCCCCGGGTGATTGATCGTGTGAACTGGAGGGCGATGTACCGGACCAATTGGCTGCGCGCAAGGTGTAGATATTCC  1607987
PD:   1196300  CAGGCGCGACAAGATGAGCTATGGGCGGTACCTGCTGCCCCGGGTGATTGATCGTGTGAACTGGAGGGCGATGTACCGGACCAATTGGCTGCGCGCAAGGTGGTGGATATTCC......  1196180

CVC:
PD:   1196181  .................................................... giPD (genomic island) 15,778 bp ..........  1180403

CVC:  1607988  -----AGCGACCGACATGACGCGGGAATGGTCACGTTGCAGACTTGCATGCAGCCGATGTATCGATTGGAACAGGCATTGAACGTGCGCGCAAGGTACGCGAT  1608103
PD:   1180402  ATTCCGGCGCCACCGACATGACGCGGGAATGGTCACGTTGCAGACTTGCATGCAGCCGATGTATCGATTGGAACAGGCATTGAACGTGCGCGCAAGGTACGCTGAT  1180282

CVC:  1608104  GCCTTGCCTGGCACGGTTGTACGGGGGGCGGTGTCTGTTTATCAATGTGCATGGACAAGACCCGTCTTTGCCGTTGATCCGGCCTCGGTCATGCCGGGGAGCAGGCTATCGCTGACG  1608224
PD:   1180283  GCGCTGCCTGGGCCCGGTGTACGGGGTGCCCGGTCGTGTTTATCAATGTGCATGGACAAGACCCGTCTTTGCCGTTGATCCGGCCTCGGTCATGCCGGGGACCAGGCTATCGCTGACG  1180163
```

B) Left and right junctions of the giCVC in   CVC 9a5c and corresponding region in   PD Temecula

```
CVC:  1638517  TCTTTCTTGCAATCGTATATGGAGCTGCATCGTAGGCATAAGAAACGAGTAGCCCTTCAGTTGGATTGGAGACAGTAGCCTATCCGTGAATGTGTCTCAAAGGCTTGGCAAATCCGAAAT  1638636
PD:   1279246  TCTTTCTTGCAATGCCATATGGAGCTGCATCGCAGGCATAGAGAACAGTAGCCTATCCGTGAATGTGTCTCAAAGGCTTGGCAAATCCTAAAT  1279127

                                                            tRNA-Gly

CVC:  1638637  GTGGTCTACAAA------------AATAGCTCAGTTGGTAGAGCGCAACCTTGCCAAGGTTGAGGTCGGAGCCTTGTTTCCGCTCCA...........  1638727
PD:   1279126  GTGGTCTACAATACCTAACACACGTGCGGAATAGCTCAGTTGGTAGAGCGCAACCTTGCCAAGGTTGAGGTCGGAGCCTTGTTTCCGCTCCA  1279018

CVC:  1638728  ................................................ giCVC (genomic island) 66,562 bp ..........  1638837
PD:

CVC:  1705399  CGTTTCCGCTCCAATGCATCCATACAATATTTTCTCAATATTTGATTTATTGATTGGCCGCAGCTGGCCCATGTAAGCGTCAGTAAGCGTCAGTAAGCGTCAGTAAGCCTGAACAGCATCGGTCAT  1705458
PD:   1279017  ATACAATCAATAGGTTGCGGATTTCTCAATATTTGATTTATTATTAATTGGCCGCAGCTTGCCCATATAGGGCGTCAATA--------------CGCGACGACATCGGCCAT  1278921

CVC:  1705459  A--TATTAGGGCGGGTACTTGGCGTAAACGTCCGTAAATCCCACTATATATGCACGATGCCCAAGCATCCTTGAATGTTCTCCGCAGTCATGCCGCTCCGAGCGGTAGCTCGGTAGCTATATCG  1705575
PD:   1278920  AATTATTAGGGCGGGTGCTTGGCGTAAACGTCCGTAAATCCCACTATATATGCACGATGCCCAAGCATCCTTGAATGTTCTCCGCAGCCATGCCCCGAGCGGTAGCTCGGTAGCCATATCG  1278801
```

FIG. 3. Comparative analysis of giPD1 and giCVC1 insertion sites. Shaded nucleotides represent the target duplication sites for each genomic island insertion. Underlined bases indicate substitutions or indels. Numbers on both sides indicate the coordinates of each genome. The coordinates for the primers used in the PCR analysis of the islands are as follows: for giPD1—P1 (PD, 1179403 to 1179420; CVC, 1608980 to 1608963), P2 (PD, 1180466 to 1180449; CVC, 1607835 to 1607852); P3 (PD, 1194838 to 1194855), and P4 (PD, 1196338 to 1196321; CVC, 1607041 to 1607058), XFCVC01R (CVC, 1639020 to 1639003), XFCVC01F (CVC, 1707041 to 1707058), and XFUN03 (PD, 1275650 to 1275667; CVC, 1707862 to 1707845).

TABLE 3. Distribution of genomic islands in different strains of *X. fastidiosa*[a]

Host	Strain	Origin	Distribution of: giCVC1 (kb)[a]	giPD1
Grapevine	Temecula	United States	− (4.5)	+
Grapevine	Fetzer	United States	− (4.5)	+
Grapevine	Traver	United States	− (4.5)	+
Grapevine	IAPAR 9715	United States	− (4.5)	+
Grapevine	ATCC 35881	United States	− (4.5)	−
Grapevine	ATCC 35877	United States	− (4.5)	−
Mulberry	ATCC 35868	United States	− (4.5)	−
Mulberry	ATCC 35869	United States	− (4.5)	−
Oleander	ATCC 700598 (Ann1)	United States	− (4.5)	−
Almond	ATCC 35870	United States	− (4.5)	+
Almond	ATCC 700965 (Dixon)	United States	− (2.9)	−
Elm	ATCC 35873	United States	− (2.9)	−
Red oak	ATCC 35874	United States	− (2.9)	ND
Plum	ATCC 35871	United States	− (2.9)	−
Plum	IAPAR 9746	Brazil	− (2.9)	−
Periwinkle	ATCC 35878	United States	− (2.9)	−
Periwinkle	IAPAR 12319	Brazil	+	−
Hibiscus	Hib#6-CCSM-BR	Brazil	+	−
Coffee	Inst. Biológico	Brazil	+	−
Coffee	C13-CCSM-BR	Brazil	+	−
Coffee	CCT 6755	Brazil	+	−
Citrus	9a5c	Brazil	+	−
Citrus	CCT 5671	Brazil	+	−
Citrus	20.06/HC16	Brazil	+	−
Citrus	Rep15	Brazil	+	−
Citrus	5R1	Brazil	+	−
Citrus	X1	Brazil	+	−
Citrus	IAPAR 11834	Brazil	+	+
Citrus	IAPAR 11775	Brazil	+	+
Citrus	IAPAR 11066	Brazil	+	+

[a] In the absence of giCVC1, the size of PCR fragments obtained with primers flanking giCVC1 is indicated. −, absence of the genomic island; +, presence of the genomic island; ND, not determined. The same results were also obtained for another 4 coffee strains and 30 citrus strains. For a full list of strains and sources, refer to the supplementary information mentioned in the text.

prophage regions are excluded and rearrangements are reoriented, the genomes of both *X. fastidiosa* strains are very similar and colinear. We propose that the evolutionary divergence of the two sequenced *X. fastidiosa* strains is thus mainly due to lateral gene transfer mediated mostly by phage vectors. It is noteworthy, however, that there have been fewer lateral gene transfer events in the *X. fastidiosa* genomes than can be detected based on a comparison with *E. coli* strains (27). Despite the genome rearrangements, the most significant conclusion to be drawn from the sequencing of the PD *X. fastidiosa* Temecula genome is that many of the genes in the two *X. fastidiosa* strains are highly similar, including not only those involved in basic cellular housekeeping but also many of those likely to have a direct role in pathogenicity. This conclusion suggests that the diseases caused by different *X. fastidiosa* pathotypes most likely rely on the expression of a common set of genes to allow the bacteria to become established in planta. This possibility of common pathogenic mechanisms implies that functional genomics studies of the two organisms would share significant common ground, and their integration might accelerate advances in combating both PD and CVC. In this regard, critical cross-infection experiments with PD *X. fastidiosa* and CVC *X. fastidiosa* strains and reciprocal hosts would be of great immediate interest to evaluate this hypothesis.

ADDENDUM IN PROOF

The polygalacturonase gene was amplified and sequenced from 11 strains of *X. fastidiosa* (see list in supplementary material). Besides the CVC *X. fastidiosa* 9a5c strain all the citrus and coffee strains examined showed the same frameshift, while other strains including mulberry, almond, and grape isolates did not show the frameshift.

REFERENCES

1. Alm, R. A., L.-S. L. Ling, D. T. Moir, B. L. King, E. D. Brown, P. C. Doig, D. R. Smith, B. Noonan, B. C. Guild, B. L. de Jonge, G. Carmel, P. J. Tummino, A. Caruso, M. Uria-Nickelsen, D. M. Mills, C. Ives, R. Gibson, D. Merberg, S. D. Mills, Q. Jiang, D. E. Taylor, G. F. Vovis, and T. J. Trust. 1999. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. Nature **397:**176–180.
2. Almeida, R. P. P., E. F. Pereira, and A. H. Purcell. 2001. Multiplication and movement of a citrus strain of *Xylella fastidiosa* within sweet orange. Plant Dis. **85:**382–386.
3. Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
4. Beretta, M. J. G., and R. P. Leite, Jr. 2000. Variegated chlorosis, p. 50–51. In L. W. Timmer, S. M. Garnsey, and J. H. Graham (ed.), Compendium of citrus diseases, 2nd ed. American Phytopathological Society, Saint Paul, Minn.
5. Borodovsky, M., and J. McIninch. 1993. GeneMark: parallel gene recognition for both DNA strands. Comput. Chem. **17:**123–133.
6. Chen, J., D. Banks, R. L. Jarret, C. J. Chang, and B. J. Smith. 2000. Use of 16S rDNA sequences as signature characters to identify *Xylella fastidiosa*. Curr. Microbiol. **40:**29–33.
7. Chen, J., R. L. Jarret, X. Qin, J. S. Hartung, D. Banks, C. J. Chang, and D. L. Hopkins. 2000. 16S rDNA sequences analysis of *Xylella fastidiosa* strains. Syst. Appl. Microbiol. **23:**349–354.
8. Costa, P. I., C. F. Franco, V. S. Miranda, D. C. Teixeira, and J. S. Hartung. 2000. Strains of *Xylella fastidiosa* rapidly distinguished by arbitrarily primed-PCR. Curr. Microbiol. **40:**279–282.
9. da Silva, A. C., J. A. Ferro, F. C. Reinach, C. S. Farah, L. R. Furlan, R. B. Quaggio, C. B. Monteiro-Vitorello, M. A. Van Sluys, N. F. Almeida, L. M. Alves, A. M. do Amaral, M. C. Bertolini, L. E. A. Camargo, G. Camarotte, F. Cannavan, J. Cardozo, F. Chambergo, L. P. Ciapina, R. M. B. Cicarelli, L. L. Coutinho, J. R. Cursino-Santos, H. El-Dorry, J. B. Faria, A. J. S. Ferreira, R. C. C. Ferreira, M. I. T. Ferro, E. F. Formighieri, M. C. Franco, C. C. Greggio, A. Gruber, A. M. Katsuyama, L. T. Kishi, R. P. Leite, Jr., E. G. M. Lemos, M. V. F. Lemos, E. C. Locali, M. A. Machado, A. M. B. N. Madeira, N. M. Martinez-Rossi, E. C. Martins, J. Meidanis, C. F. M. Menck, C. Y. Miyaki, D. H. Moon, L. M. Moreira, M. T. M. Novo, V. K. Okura, M. C. Oliveira, V. R. Oliveira, H. A. Pereira, Jr., A. Rossi, J. A. D. Sena, C. Silva, R. F. de Souza, L. A. F. Spinola, M. A. Takita, R. E. Tamura, E. C. Teixeira, R. I. D. Tezza, M. Trindade dos Santos, D. Truffi, S. M. Tsai, F. F. White, J. C. Setubal, and J. P. Kitajima. 2002. Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. Nature **417:**459–463.
10. Delcher, A. L., D. Harmon, S. Kasif, O. White, and S. L. Salzberg. 1999. Improved microbial gene identification with GLIMMER (306K, PDF format). Nucleic Acids Res. **27:**4636–4641.
11. Delcher, A. L., S. Kasif, R. D. Fleischmann, J. Peterson, O. White, and S. L. Salzberg. 1999. Alignment of whole genomes. Nucleic Acids Res. **27:**2369–2376.
12. Dow, J. M., A. E. Osbourn, T. J. Wilson, and M. J. Daniels. 1995. A locus determining pathogenicity of *Xanthomonas campestris* is involved in lipopolysaccharide biosynthesis. Mol. Plant-Microbe Interact. **8:**768–777.
13. Fleischmann, R. D., M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J.-F. Tomb, B. A. Dougherty, J. M. Merrick, K. McKenney, G. G. Sutton, W. FitzHugh, C. A. Fields, J. D. Gocayne, J. D. Scott, R. Shirley, L. I. Liu, A. Glodek, J. M. Kelley, J. F. Weidman, C. A. Phillips, T. Spriggs, E. Hedblom, M. D. Cotton, T. Utterback, M. C. Hanna,

D. T. Nguyen, D. M. Saudek, R. C. Brandon, L. D. Fine, J. L. Fritchman, J. L. Fuhrmann, N. S. Geoghagen, C. L. Gnehm, L. A. McDonald, K. V. Small, C. M. Fraser, H. O. Smith, and J. C. Venter. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science **269**:496–512.

14. Freitag, J. H. 1951. Host range of Pierce's disease virus of grapes as determined by insect transmission. Phytopathology **41**:920–934.

15. Gordon, D., C. Abajian, and P. Green. 1998. Consed: a graphical tool for sequence finishing. Genome Res. **8**:195–202.

16. Hendson, M., A. H. Purcell, D. Chen, C. Smart, M. Guilhabert, and B. Kirkpatrick. 2001. Genetic diversity of Pierce's disease strains and other pathotypes of *Xylella fastidiosa*. Appl. Environ. Microbiol. **67**:895–903.

17. Hopkins, D. L. 1989. *Xylella fastidiosa*: a xylem-limited bacterial pathogen of plants. Annu. Rev. Phytopathol. **27**:271–290.

18. Hughes, D. 2000. Evaluating genome dynamics: the constraints on rearrengements within bacterial genomes. Genome Biol. **6**:1–8.

19. Kamper, S. M., W. J. French, and S. R. deKloet. 1985. Genetic relationships of some fastidious xylem-limited bacteria. Int. J. Syst. Bacteriol. **35**:185–188.

20. Kholodii, G. Ya., O. V. Yurieva, O. L. Lomovskaya, Zh. M. Gorlenko, S. Z. Mindlin, and V. G. Nikiforov. 1993. Tn5053, a mercury resistance transposon with integron's ends. J. Mol. Biol. **230**:1103–1107.

21. Lowe, T. M., and S. R. Eddy. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequences. Nucleic Acids Res. **25**:955–964.

22. McClelland, M., K. E. Sanderson, J. Spieth, S. W. Clifton, P. Latreille, L. Courtney, S. Porwollik, J. Ali, M. Dante, F. Du, S. Hou, D. Layman, S. Leonard, C. Nguyen, K. Scott, A. Holmes, N. Grewal, E. Mulvaney, E. Ryan, H. Sun, L. Florea, W. Miller, T. Stoneking, M. Nhan, R. Waterston, and R. K. Wilson. 2001. Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. Nature **413**:852–856.

23. McDaniel, T. K., and J. B. Kaper. 1997. A cloned pathogenicity island from enteropathogenic *Escherichia coli* confers the attaching and effacing phenotype on *E. coli* K-12. Mol. Microbiol. **23**:399–407.

24. Mehta, A., and Y. R. Rosato. 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.

25. Myers, E. W., and W. Miller. 1988. Optimal alignments in linear space. Comput. Appl. Biosci. **4**:11–17.

26. Ochman, H., J. G. Lawrence, and E. A. Groisman. 2000. Lateral gene transfer and the nature of bacterial innovation. Nature **405**:299–304.

27. Perna, N. T., G. Plunkett, I. I. I., V. Burland, B. Mau, J. D. Glasner, D. J. Rose, G. F. Mayhew, P. S. Evans, J. Gregor, H. A. Kirkpatrick, G. Posfai, J. Hackett, S. Klink, A. Boutin, Y. Shao, L. Miller, E. J. Grotbeck, N. W., Davis, A. Lim, E. Dimalanta, K. Potamousis, J. Apodaca, T. S. Anantharaman, J. Lin, G. Yen, D. C. Schwartz, R. A. Welch, and F. R. Blattner. 2001. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. Nature **409**:529–533.

28. Pooler, M. R., and J. S. Hartung. 1995. Specific PCR detection and identification of *Xylella fastidiosa* strains causing citrus variegated chlorosis. Curr. Microbiol. **31**:377–381.

29. Purcell, A. H., and D. H. Hopkins. 1996. Fastidious xylem-limited bacterial plant pathogens. Annu. Rev. Phytopathol. **34**:131–151.

30. Qin, X., V. S. Miranda, M. A. Machado, E. G. M. Lemos, and J. S. Hartung. 2001. An evaluation of the genetic diversity of *Xylella fastidiosa* isolated from diseased citrus and coffee in São Paulo, Brazil. J. Bacteriol. **91**:599–605.

31. Ravatn, R., S. Studer, A. J. Zehnder, and J. R. van der Meer. 1998. Int-B13,

an unusual site-specific recombinase of the bacteriophage P4 integrase family, is responsible for chromosomal insertion of the 105-kilobase *clc* element of *Pseudomonas* sp. strain B13. J. Bacteriol. **180**:5505–5514.

32. Salanoubat, M., S. Genin, F. Artiguenave, J. Gouzy, S. Mangenot, M. Arlat, A. Billault, P. Brottier, J. C. Camus, L. Cattolico, M. Chandler, N. Choisne, C. Claudel-Renard, S. Cunnac, N. Demange, C. Gaspin, M. Lavie, A. Moisan, C. Robert, W. Saurin, T. Schiex, P. Siguier, P. Thebault, M. Whalen, P. Wincker, M. Levy, J. Weissenbach, and C. A. Boucher. 2002. Genome sequence of the plant pathogen *Ralstonia solanacearum*. Nature **415**:497–502.

33. Simpson, A. J. G., F. C. Reinach, P. Arruda, F. A. Abreu, M. Acencio, R. Alvarenga, L. M. C. Alves, J. E. Araya, G. S. Baia, C. S. Baptista, M. H. Barros, E. D. Bonaccorsi, S. Bordin, J. M. Bové, M. R. S. Briones, M. R. P. Bueno, A. A. Camargo, L. E. A. Camargo, D. M. Carraro, H. Carrer, N. B. Colauto, C. Colombo, F. F. Costa, M. C. R. Costa, C. M. Costa-Neto, L. L. Coutinho, M. Cristofani, E. Dias-Neto, C. Docena, H. El-Dorry, A. P. Facincani, A. J. S. Ferreira, V. C. A. Ferreira, J. A. Ferro, J. S. Fraga, S. C. França, M. C. Franco, M. Frohme, L. R. Furlan, M. Garnier, G. H. Goldman, M. H. S. Goldman, S. L. Gomes, A. Gruber, P. L. Ho, J. D. Hoheisel, M. L. Junqueira, E. L. Kemper, J. P. Kitajima, J. E. Krieger, E. E. Kuramae, F. Laigret, M. R. Lambais, L. C. C. Leite, E. G. M. Lemos, M. V. F. Lemos, S. A. Lopes, C. R. Lopes, J. A. Machado, C. A. Machado, A. M. B. N. Madeira, H. M. F. Madeira, C. L. Marino, M. V. Marques, E. A. L. Martins, E. M. F. Martins, A. Y. Matsukuma, C. F. M. Menck, E. C. Miracca, C. Y. Miyaki, C. B. Monteiro-Vitorello, D. H. Moon, M. A. Nagai, A. L. T. O. Nascimento, L. E. S. Netto, A. Nhani, Jr., F. G. Nobrega, L. R. Nunes, M. A. Oliveira, M. C. de Oliveira, R. C. de Oliveira, D. A. Palmieri, A. Paris, B. R. Peixoto, G. A. G. Pereira, H. A. Pereira, Jr., J. B. Pesquero, R. B. Quaggio, P. G. Roberto, V. Rodrigues, J. de M. Rosa, V. E. de Rosa, Jr., R. G. de Sá, R. V. Santelli, H. E. Sawasaki, A. C. R. da Silva, A. M. da Silva, F. R. da Silva, W. A. Silva, Jr., J. F. da Silveira, M. L. Z. Silvestri, W. J. Siqueira, A. A. de Souza, A. P. de Souza, M. F. Terenzi, D. Truffi, S. M. Tsai, M. H. Tsuhako, H. Vallada, M. A. Van Sluys, S. Verjovski-Almeida, A. L. Vettore, M. A. Zago, M. Zatz, J. Meidanis, and J. C. Setubal. 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. Nature **406**:151–157.

34. Slauch, J. M., A. A. Lee, M. J. Mahan, and J. J. Mekalanos. 1996. Molecular characterization of the *oafA* locus responsible for acetylation of *Salmonella typhimurium* O antigen: OafA is a member of a family of integral membrane *trans*-acylases. J. Bacteriol. **178**:5904–5909.

35. Sullivan, J. T., and C. W. Robson. 1998. Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. Proc. Natl. Acad. Sci. USA **95**:5145–5149.

36. Van Sluys, M. A., C. B. Monteiro-Vitorello, L. E. A. Camargo, C. F. M. Menck, A. C. R. da Silva, J. A. Ferro, M. C. Oliveira, J. C. Setubal, J. P. Kitajima, and A. J. Simpson. 2002. Comparative genome analysis of plant-associated bacteria. Annu. Rev. Phytopathol. **40**:09.1–09.21.

37. Varela, L. G., R. J. Smith, and P. A. Phillips. 2001. Pierce's disease. University of California Agriculture and Natural Resources publication 21600. University of California.

38. Wells, J. M., B. C. Raju, H.-Y. Hung, W. G. Weisburg, L. Mandelco-Paul, and D. J. Brenner. 1987. *Xylella fastidiosa* new-genus new-species gram-negative xylem-limited fastidious plant bacteria related to *Xanthomonas* spp. Int. J. Syst. Bacteriol. **37**:135–143.

39. Williams, K. P. 2002. Integration sites for genetic elements in prokaryotic tRNA and tmRNA genes: sublocation preference of integrase subfamilies. Nucleic Acids Res. **30**:866–875.