Comparative genomic analysis of coffee-infecting Xylella fastidiosa strains isolated from Brazil

Deibs Barbosa,1 Valquiria Campos Alencar,1 Daiene Souza Santos,1 Ana Claudia de Freitas Oliveira,1 Alessandra A. de Souza,2 Helvecio D. Coletta-Filho,2 Regina Costa de Oliveira1 and Luiz R. Nunes3

Correspondence
Luiz R. Nunes
Luiz.Nunes@ufabc.edu.br

Received 12 November 2014
Accepted 28 February 2015

1Núcleo Integrado de Biotecnologia, Universidade de Mogi das Cruzes, Av. Dr Cândido Xavier de Almeida Souza, 200, Mogi das Cruzes, SP, CEP 08780-911, Brazil
2Centro APTA Citros Sylvio Moreira – Instituto Agronômico de Campinas, Caixa Postal 04, Condeirópolis, SP, CEP 13490-970, Brazil
3Centro de Ciências Naturais e Humanas, Universidade Federal do ABC, Rua Santa Adélia, 166, Santo André, SP, CEP 09210-170, Brazil

Strains of Xylella fastidiosa constitute a complex group of bacteria that develop within the xylem of many plant hosts, causing diseases of significant economic importance, such as Pierce’s disease in North American grapevines and citrus variegated chlorosis in Brazil. X. fastidiosa has also been obtained from other host plants, in direct correlation with the development of diseases, as in the case of coffee leaf scorch (CLS) – a disease with potential to cause severe economic losses to the Brazilian coffee industry. This paper describes a thorough genomic characterization of coffee-infecting X. fastidiosa strains, initially performed through a microarray-based approach, which demonstrated that CLS strains could be subdivided in two phylogenetically distinct subgroups. Whole-genomic sequencing of two of these bacteria (one from each subgroup) allowed identification of ORFs and horizontally transferred elements (HTEs) that were specific to CLS-related X. fastidiosa strains. Such analyses confirmed the size and importance of HTEs as major mediators of chromosomal evolution amongst these bacteria, and allowed identification of differences in gene content, after comparisons were made with previously sequenced X. fastidiosa strains, isolated from alternative hosts. Although direct experimentation still needs to be performed to elucidate the biological consequences associated with such differences, it was interesting to verify that CLS-related bacteria display variations in genes that produce toxins, as well as surface-related factors (such as fimbrial adhesins and LPS) that have been shown to be involved with recognition of specific host factors in different pathogenic bacteria.

INTRODUCTION

Xylella fastidiosa is a Gram-negative bacterium that develops within the xylem of many plant hosts, such as plums, mulberries, pears, oranges, grapes, coffee, almond

1These authors contributed equally to this work.

Abbreviations: ANI, average nucleotide identity; CLS, coffee leaf scorch; CVC, citrus variegated chlorosis; HTE, horizontally transferred element; MCMC, Monte Carlo Markov chain.

Original microarray data have been deposited in the NCBI Gene Expression Omnibus under accession number GSE41163 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE41163). Whole-genome shotgun projects for strains 32 and 6c have been deposited at DDBJ/EMBL/GenBank under the accession numbers AWYH00000000 and AXBS00000000, respectively.

Supplementary material is available with the online Supplementary Material.
made *X. fastidiosa* a major agronomical concern in the American continent and this micro-organism has been the subject of increasing attention by many research programs since the mid-1990s (Doddapaneni et al., 2006, 2007). As a consequence of these efforts, one *X. fastidiosa* strain (9a5c), associated with CVC in Brazil, was the first plant bacterium to have its complete genome elucidated (Simpson et al., 2000). Sequencing efforts have been extended to other *X. fastidiosa* strains and, currently, draft or complete genomic sequences have been obtained for a variety of strains, obtained from different hosts, including almond trees (strains M12, M23 and Dixon), oleander (strain Ann-1), grapevine (strains Temecula1 and GB514), elderberry (strain EB92-1), oak (strain Griffin-1), mulberry (strain Mul-MD), sycamore (strain Sy-VA) and pear trees (strain PLS229) (Zhang et al., 2011; Varani et al., 2012; Chen et al., 2013; Guan et al., 2014a, b; Su et al., 2014).

Genomic sequences from these *X. fastidiosa* strains have been submitted to extensive in silico evaluations, assisting researchers to identify and categorize genome-wide DNA variations, as well as their influence on strain functional divergence (Doddapaneni et al., 2006; Varani et al., 2012). Multiple alignment of chromosomal sequences identified single nucleotide polymorphisms and indels that can now be used to estimate the relative similarity between the strains and their rates of genomic evolution, which seem to be different for each of them (Doddapaneni et al., 2006). All ORFs unique to each strain have been catalogued, and their sequences have been used to design primer pairs and probes to assist in PCR-based detection of different *X. fastidiosa* strains in the wild (Doddapaneni et al., 2006, 2007).

Additionally, these comparisons have shown that the bacterial species *X. fastidiosa*, originally characterized from 25 isolates (obtained from 10 different hosts), constitutes a significantly complex group of plant-associated bacteria (Schaad et al., 2004; Almeida et al., 2008; Nunney et al., 2010). All strains appear to share a common set of genes, which constitute the core gene pool, including most genes required for maintenance of bacterial metabolism and replication. However, all strains sequenced so far carry a large flexible gene pool, which accounts for ~18% of the total genome size (Nunes et al., 2003), constituted of integrated prophages, plasmids, genomic islands and other horizontally transferred elements (HTEs), which have been shown to carry genes that may assist different strains to adapt to their specific niches (Ochman et al., 2000; Gal-Mor & Finlay, 2006; Kittichotirat et al., 2011; Fernández-Gómez et al., 2012; Acuña et al., 2013).

More importantly, comparative studies have shown that a surprisingly large proportion of genes present in the flexible gene pool of *X. fastidiosa* are shared amongst distantly related strains, indicating the possibility that the bacteria are actively and constantly exchanging genetic material in the wild (Nunes et al., 2003; da Silva et al., 2007) – a possibility that is reinforced by the recent demonstration that *X. fastidiosa* cells display natural competence for transformation via uptake of exogenous DNA (Kung & Almeida, 2011; Kung et al., 2013). Comparative analyses also suggest that *X. fastidiosa* genomes might evolve at a very fast pace, taking advantage of such a large number of HTEs and the fact that different strains are constantly coming into contact with one another through mixed infections that occur in the same host (particularly in insects) (Chen et al., 2005; Costa et al., 2006). Genetic exchange amongst these bacteria may contribute not only to variability, but also to the development of distinct infective/virulence capacities in different strains. For example, a recent study (Nunney et al., 2012) suggested that South American *X. fastidiosa* strains might have become pathogenic to citrus plants after gaining genetic variation via intrasubspecific recombination, which may have facilitated a switch from their native hosts. Thus, genomic characterization of other *X. fastidiosa* strains, particularly those known to coexist in the same geographical area, is vital in order to understand the evolution of this important group of phytopathogens (Chen et al., 2005).

However, most of the genomic information available for *X. fastidiosa* has been obtained from North American strains (Varani et al., 2013) and the relative lack of sequence information from South American strains may have introduced biases to some of the conclusions drawn from recent genomic studies within the group (Doddapaneni et al., 2006, 2007). So far, there is only one work that attempted to perform a genomic survey of South American *X. fastidiosa* strains (da Silva et al., 2007) and this study did not involve whole-genomic sequencing, but microarray-based comparisons, followed by suppressive subtractive hybridization. Moreover, these analyses focused only on citrus-infecting *X. fastidiosa* strains, overlooking the contribution of coffee-related strains to the evolutionary history of the South American *X. fastidiosa* group. Coffee-infecting strains are believed to be responsible for another plant disease in South America, known as coffee leaf scorch (CLS), which has the potential to cause severe economic losses. CLS symptoms include drying of infected branches, shortening of internode regions, chlorosis and early senescence of leaves, as well as decreased fruit size, affecting overall plant productivity (de Lima et al., 1998). The disease was originally described in 1997, in orchards distributed across the states of São Paulo and Minas Gerais in Brazil (Paradela-Filho et al., 1997), but some authors argue that such symptoms had already been observed earlier, although mistakenly attributed to nematode infections (Li et al., 2001). More recently, the disease has been identified in other important coffee-growing regions of Brazil throughout the states of Espírito Santo, Paraná, Bahia and Goiás (Rocha et al., 2010).

All comparative analyses performed with CVC- and CLS-related *X. fastidiosa* strains suggest that they are extremely similar, sharing >99% sequence similarity amongst classic molecular markers, such as the 16S rRNA gene and the 16S–23S intergenic spacer (Mehta & Rosato, 2001; Mehta et al., 2006). However, several genetic variations have been found amongst these strains, suggesting that *X. fastidiosa* is a heterogeneous population comprised of different ecotypes or species. Comparative genomic analysis of *X. fastidiosa* genomes has been performed with the aim of identifying these genetic variations and understanding their ecological and evolutionary significance. In this context, this paper aims to provide a comprehensive overview of the genomic diversity observed in *X. fastidiosa* strains collected from different regions of the world, with a particular emphasis on the genetic variations associated with the disease-causing potential of these bacteria.
et al., 2001). To this date, however, very little is known about the genomic structure and composition of CLS-related strains, which constitutes a significant agronomical concern, given the potential damage that this disease can cause to the Brazilian coffee industry, and as CVC- and CLS-associated strains, although biologically distinct, have been suggested to undergo frequent recombination in the wild (Almeida et al., 2008; Nunney et al., 2012).

Here, we describe the use of microarray hybridization analyses to perform a preliminary characterization of the genomic structure in different X. fastidiosa strains, obtained from infected coffee plants, distributed across the states of São Paulo and Minas Gerais in Brazil. These analyses suggested that chromosomal structure and gene content amongst CLS-related bacteria are in accord with their subdivision into two previously described, phylogenetic distinct, subgroups (Almeida et al., 2008; Nunney et al., 2012). Draft genomes obtained for two of these strains (each representing a different subgroup) were submitted to comparative analyses with the fully sequenced genomes obtained from other X. fastidiosa strains. The resulting data unravel characteristics that help to explain the evolution of South American X. fastidiosa strains, indicating significant structural differences amongst the genomes of this important group of phytopathogens.

METHODS

Strains, growth conditions and DNA extraction. All X. fastidiosa strains used in this work (Table 1) are deposited at the culture collections of both the Centro APTA de Citros Sylvio Moreira and the Laboratory of Insect Vectors of Phytopathogens (Department of Entomology and Aecology, ESALQ/Universidade de São Paulo, Brazil). Those interested in obtaining more specific information should contact H. D. C. F. (helvecio@centrodecitricicultura.br) or João R. S. Lopes (jrslopes@usp.br).

Primary isolation of the strains was performed as described previously (da Silva et al., 2007). Briefly, symptomatic leaves from infected plants were collected, washed and disinfected; the petioles were placed in a plastic tube with PBS and ground in a tissue grinder. Suspensions were collected, washed and disinfected; the petioles were placed in a plastic tube with PBS and ground in a tissue grinder. Suspensions were diluted 10- and 100-fold, plated onto PW agar plates (Davis et al., 1950) and incubated at 28 °C for 15–20 days. Isolated colonies were selected using a stereo microscope, streaked onto fresh PW agar plates and incubated at 28 °C for 7 days. The latter procedure was repeated twice for purification of each strain. Identification of X. fastidiosa was carried out based on fastidious in vitro growth, white colour of colonies and PCR assays using specific primers (Pooler & Hartung, 1995). All bacteria were maintained at ~80 °C. For this work, the bacteria were recovered on PW agar medium and the plates maintained for 10 days at 28 °C. The colonies were transferred once to new plates containing the same medium, grown for 20 days and harvested for DNA extraction, as described previously (da Silva et al., 2007).

Microarray hybridization analyses. X. fastidiosa DNA microarrays were constructed as described previously (Costa de Oliveira et al., 2002). DNA labelling, hybridization conditions and overall data analyses were performed as described by Nunes et al. (2003) and da Silva et al. (2007). ORFs whose spots in the array showed a mean reference/test ratio >5:1 were considered to be missing in the test strain. The application of these criteria in a direct sequence comparison between X. fastidiosa strains 9a5c and Temecula1, which have been sequenced completely, provided an estimated error rate <0.3 % (Costa de Oliveira et al., 2002). Raw and normalized data from all microarray hybridizations, as well as the microarray complete annotation file, have been deposited, in MIAME-compliant format, in the NCBI Gene Expression Omnibus (GSE41163).

Genomic sequencing and comparative analyses of the coffee-related X. fastidiosa strains. DNA samples from strains 6c and 32 (10 µg each) were used to construct shotgun libraries with the aid of a GS DNA Library Preparation kit (Life Sciences/Roche). For mate pair library construction, total genomic DNA was randomly sheared into 8 kb average fragments using a Covaris S-series system. Whole-genome sequencing was performed in a Genome Sequencer FLX 454, using reagents from the GS FLX reagents, according to the manufacturer’s protocols and instructions (Life Sciences/Roche), and as described by Margulies et al. (2005).

Sequence analysis and assembly were performed with GS FLX version 2.5.3 and GS De novo Assembler version 2.5.3, using the default parameters established by Life Sciences/Roche. Scaffolds derived from the de novo assembly of each bacterium were ordered using the genome of strain 9a5c as a reference, with the aid of Mauve software, using the progressive alignment algorithm (Darling et al., 2010). Draft genomes of strains 6c and 32 were annotated through submission to the NCBI Prokaryotic Genome Automatic Annotation Pipeline. The whole-genome shotgun projects have been deposited in GenBank under the accession numbers AWYH10000000 (strain 32) and AXXBS00000000 (strain 6c). Mauve was also used to identify genes exclusive to each strain and/or in groups of strains, allowing categorization of homologous ORFs in a pangenome for the four X. fastidiosa strains.

Prophage regions were identified in the draft genomes of the CLS-related strains using the online PHAe Search Tool (PHAST) developed by Zhou et al. (2011). Average nucleotide identity (ANI) values amongst the available X. fastidiosa genomes were estimated with the aid of MCMC generation burn-in and a 16-category Carlo Markov chains (MCMCs), 0.5 annealing temperature, 100 000 MCMC generation burn-in and a 16-category distribution. A
A consensus tree was generated after burn-in, using a 50% majority rule, obtained from 75 cladograms. Credibility values, representing the posterior probability of each clade, were generated automatically by the software. Rooting of the tree was established using the information regarding gene content from two Xanthomonas species as outgroup. The final tree was then edited with the online application iTOL (Letunic & Bork, 2007).

RESULTS

Microarray-based genomic characterization of CLS-associated X. fastidiosa strains

In an attempt to obtain an instant overview of the genomic structure of CLS-associated X. fastidiosa strains, we conducted microarray-based comparisons amongst eight different strains, using as reference a biochip containing the genome of the citrus-associated strain 9a5c, which has been used previously to conduct genomic comparisons amongst different strains of this micro-organism (Nunes et al., 2003; da Silva et al., 2007). All coffee-related X. fastidiosa strains came from infected coffee trees distributed across different regions of the states of São Paulo and Minas Gerais, Brazil. As observed in Fig. 1, the genomic profiles obtained for these eight coffee-related strains suggested their subdivision into two major groups, mostly characterized by differences that mapped into putative mobile elements, previously identified in the genome of strain 9a5c (Nunes et al., 2003; Chen et al., 2005). The first subgroup (group I), containing strains 9c, 32, iso32 and 33, was characterized by deletions that mapped into two prophage elements of strain 9a5c: \( XfP4 \) and \( XfP6 \) (Simpson et al., 2000; de Mello Varani et al., 2008). The later prophage seems to be completely absent from the genome of these bacteria, whilst the former seems to be incomplete, as a series of deletions across its structure could be observed in the genomes of the above-mentioned CLS strains. A second subgroup (group II) was composed of strains 08, 6c, 48 and 23, in which elements from prophages \( XfP4 \) and \( XfP6 \) (de Mello Varani et al., 2008) were mostly present, whereas ORFs from a plasmid-derived insertion element (PI1) (Nunes et al., 2003) were absent, suggesting that this plasmid-mediated insertion might not have occurred in this particular subgroup of coffee-associated strains. Moreover, this later group carried sequences similar to \( pXf51 \), the large 51 kb plasmid found in strain 9a5c (Simpson et al., 2000).

Table 1. X. fastidiosa strains used in the current study

Microarray data regarding the coffee-infecting strains are described for the first time in the current paper. Data regarding the remaining strains were presented previously by Nunes et al. (2003) and da Silva et al. (2007).

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Host</th>
<th>Locality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Almond*</td>
<td>Prunus amygdalus</td>
<td>California, USA</td>
</tr>
<tr>
<td>Citrus 187b†</td>
<td>Citrus sinensis</td>
<td>Ubarana, SP, Brazil</td>
</tr>
<tr>
<td>Citrus 36f†</td>
<td>Citrus sinensis</td>
<td>Ubarana, SP, Brazil</td>
</tr>
<tr>
<td>Citrus 56a†</td>
<td>Citrus sinensis</td>
<td>Ubarana, SP, Brazil</td>
</tr>
<tr>
<td>Citrus 912c†</td>
<td>Citrus sinensis</td>
<td>Gavião Peixoto, SP, Brazil</td>
</tr>
<tr>
<td>Citrus Cv21†</td>
<td>Citrus sinensis</td>
<td>Colina, SP, Brazil</td>
</tr>
<tr>
<td>Citrus Fb7†</td>
<td>Citrus sinensis</td>
<td>Corrientes, Argentina</td>
</tr>
<tr>
<td>Citrus SJ*</td>
<td>Citrus sinensis</td>
<td>Taquaritinga, SP, Brazil</td>
</tr>
<tr>
<td>Coffee 08‡</td>
<td>Coffea arabica</td>
<td>Pedregulho, SP, Brazil</td>
</tr>
<tr>
<td>Coffee 13*</td>
<td>Coffea arabica</td>
<td>Cordeirópolis, SP, Brazil</td>
</tr>
<tr>
<td>Coffee 23‡</td>
<td>Coffea arabica</td>
<td>Garça, SP, Brazil</td>
</tr>
<tr>
<td>Coffee 32‡</td>
<td>Coffea arabica</td>
<td>Muritinga do Sul, SP, Brazil</td>
</tr>
<tr>
<td>Coffee 33‡</td>
<td>Coffea arabica</td>
<td>Varginha, MG, Brazil</td>
</tr>
<tr>
<td>Coffee 48‡</td>
<td>Coffea arabica</td>
<td>Lavras, MG, Brazil</td>
</tr>
<tr>
<td>Coffee 6c‡</td>
<td>Coffea arabica</td>
<td>Matão, SP, Brazil</td>
</tr>
<tr>
<td>Coffee 9c‡</td>
<td>Coffea arabica</td>
<td>Garça, SP, Brazil</td>
</tr>
<tr>
<td>Coffee 9c iso32‡</td>
<td>Coffea arabica</td>
<td>Muritinga do Sul, SP, Brazil</td>
</tr>
<tr>
<td>Elm*</td>
<td>Ulmus americana</td>
<td>Washington, USA</td>
</tr>
<tr>
<td>Grape CA*</td>
<td>Vitis sp.</td>
<td>California, USA</td>
</tr>
<tr>
<td>Grape FL*</td>
<td>Vitis sp.</td>
<td>Florida, USA</td>
</tr>
<tr>
<td>Mulberry*</td>
<td>Morus rubra</td>
<td>Massachusetts, USA</td>
</tr>
<tr>
<td>Plum I*</td>
<td>Prunus salicina</td>
<td>Ponta Grossa, PR, Brazil</td>
</tr>
<tr>
<td>Plum II*</td>
<td>Prunus sp.</td>
<td>Georgia, USA</td>
</tr>
<tr>
<td>Ragweed*</td>
<td>Ambrosia artemisiifolia</td>
<td>Florida, USA</td>
</tr>
</tbody>
</table>

* X. fastidiosa isolate studied previously by Nunes et al. (2003).
† X. fastidiosa isolate studied previously by da Silva et al. (2007).
‡ X. fastidiosa isolate studied for the first time in this work.
Interestingly, the subdivision of CLS-related strains into two distinct subgroups is in accord with the most recent hypotheses regarding the evolutionary history of South American X. fastidiosa strains, as described by Almeida et al. (2008) and Nunney et al. (2012). Using MLST markers, these authors suggested that coffee-related strains appeared earlier during evolution of the South American bacteria. These earlier strains were likely to have split into two subgroups: a basal monophyletic CLS subgroup and a CLS/citrus-infecting subgroup. Thus, to check if the CLS strains characterized by our microarray hybridizations correlated with the two phylogenetically distinct subgroups defined through the MLST data, we decided to employ a microarray hybridization-based phylogenomic approach, as originally described by Dorrell et al. (2005), to reconstruct the evolutionary history of X. fastidiosa strains. In this type of approach, information regarding the presence or absence of each gene, gathered by microarray hybridization experiments, is used to generate a table of discrete characters that represent each genome to be compared (Champion et al., 2005). This table is then used to calculate the evolutionary history of the group through Bayesian statistics or some other alternative method (Howard et al., 2006). This phylogenomic approach is ideal to obtain a genome-based phylogeny for X. fastidiosa, as microarray-based comparisons have been performed with a great number of X. fastidiosa strains, including North American (Nunes et al., 2003) and South American bacteria, obtained from both infected citrus (da Silva et al., 2007) and coffee trees (Fig. 1). All strains used to perform this analysis are described in Table 1.

Thus, using the genome of strain 9a5c as a reference, we gathered information regarding the presence/absence of individual genes/ORFs in the genomes of 24 X. fastidiosa strains to reconstruct the phylogenomic tree shown in Fig. 2. This tree was reconstructed with only 1871 genes/ORFs, as elements associated with HTEs were excluded in order to prevent biasing the phylogenetic analysis with sequences that were not present in the ancestral states (Marttinen et al., 2012). Moreover, we used information regarding the presence/absence of homologues in the genomes of two Xanthomonas species in order to establish a reliable outgroup to the X. fastidiosa group [information regarding the similarities in genomic content amongst X. fastidiosa and Xanthomonas strains was reported previously by Pieretti et al. (2009) using direct in silico comparisons]. As observed in Fig. 2, this microarray-based phylogenomic analysis provided a tree that correlated with previous analyses, supporting the current subdivision of X. fastidiosa into three different subspecies (Schaad et al., 2004). However, the credibility values observed amongst the North American strains were not very high – probably due to the fact that all microarray hybridizations used a South American strain (9a5c) as a

---

**Fig. 1.** Genomic comparison of X. fastidiosa isolates through microarray hybridization analyses. The figure shows a schematic representation of ORF distribution in the genomes of the X. fastidiosa strains analysed in this study, using as a reference the genome of strain 9a5c. Chromosomal ORFs are linearly represented from ORF 0001 to 2782, followed by ORFs from plasmid pXf51 (from ORF a0001 to a0065). ORFs present in the genome of each tested isolate are represented in red, whilst missing ORFs are represented in black. The genomic profiles were evaluated by a clustering algorithm, with the aid of tmap software (Saeed et al., 2003), which resulted in their subdivision into two subgroups. The hierarchical profile derived from the clustering analysis is shown at the top of the figure. The figure also displays the localization of HTEs in the genome of strain 9a5c (see Nunes et al., 2003; de Mello Varani et al., 2008).
reference. Thus, these low credibility values indicated that the phylogenetic history of North American strains, as described in Fig. 2, must be considered with the necessary caution. However, the South American branch of the tree depicted a much more reliable pattern, characterized by a single monophyletic clade with a credibility value of 1.0. This group could be further subdivided into three monophyletic subgroups, each of them presenting a credibility value of 1.0: the two most basal groups were composed of coffee-infecting strains, whilst the latter was composed exclusively by citrus-infecting strains, emerging as a sister clade to the second CLS-related subgroup. This pattern was in full accord with the MLST analyses of Almeida et al. (2008) and Nunney et al. (2012), confirming that the chromosomal structural differences identified by our microarray hybridization analyses reflected the phylogenetic relationships amongst the CLS-related strains under study.

**Sequencing and assembly of draft genomes from CLS-related X. fastidiosa strains**

As the phylogenomic evaluation confirmed that the two subgroups of CLS-related strains represented two evolutionary distinct clades (one of them being very close to the citrus-infecting bacteria), we decided to obtain a more detailed view of the genomic structure of CLS-related X. fastidiosa strains.
**fastidiosa** by selecting two strains – one from each subgroup – and submitting them to genomic sequencing, with the aid of next-generation sequencing technology (Alencar et al., 2014). Thus, DNA samples from strains 32 (group I) and 6c (group II) were used to construct shotgun libraries that were pyrosequenced with the aid of a Roche 454 sequencer. Additionally, 8 kb mate pair libraries of the two strains were constructed and sequenced to assist in the final assembly of the contigs into scaffolds. Finally, scaffolds were ordered with the aid of the progressiveMauve algorithm (Darling et al., 2010), using the complete genome of strain 9a5c as a reference, resulting in two draft genomes of ~2.56 (strain 6c) and ~2.60 (strain 32) Mbp. The evolutionary proximity between the CLS-related and other *X. fastidiosa* strains that had their genomes sequenced was next verified through the evaluation of their ANI values (Richter & Rossello-Móra, 2009), confirming that CLS-related strains were closer to the citrus-infecting strain 9a5c than to any other North American *X. fastidiosa* (Table 2).

To obtain a more detailed comparative analysis of the draft genomes of the CLS strains, they were aligned with the fully sequenced chromosomes from strains 9a5c and Temecula1, as shown in Fig. 3. This syntenic evaluation took the chromosome of strain 9a5c as a reference and subdivided it in seven collinear blocks, labelled A–G. Note that the chromosomal structure for all *X. fastidiosa* strains could be rebuilt through shuffling and/or subdivision of these seven blocks, allowing a thorough visualization of how homologous sequences were distributed along the chromosomes from each strain (Fig. 3). As expected, the draft genomes obtained for the two CLS-associated strains displayed a very similar chromosomal structure and most differences between them were in accord with the features predicted by the microarray experiments.

**Characterization of HTEs in CLS-related X. fastidiosa strains.**

Several studies suggested that transpositional activity and recombination of HTEs might represent an important driving force of chromosomal evolution amongst *X. fastidiosa* strains (Nunes et al., 2003; Van Sluys et al., 2003; de Mello Varani et al., 2008, Varani et al., 2013). Thus, we decided to verify the presence of such elements in the draft genomes of the two CLS-associated strains. This analysis was made with the aid of PHAST, and resulted in the identification of six prophages and/or phage remnants distributed along the genomes of both strains 6c and 32 (Fig. 3). Moreover, the Mauve analysis indicated that many of these elements seemed to correlate with prophages already mapped in the genome of strains 9a5c and Temecula1, whilst others represented new elements, which had not been previously identified in the currently known *X. fastidiosa* pangenome. In order to produce a more detailed comparison of the prophage elements that could be found within the South American *X. fastidiosa* subgroup, we aligned all prophages found in the genomes of strains 9a5c, 6c and 32, subdividing them in seven groups of corresponding elements, as visualized in Fig. 4.

This analysis indicated that the three strains have four prophages in common: XfP3, XfP4, XfP5 and XfP1/XfP2 (which was present as a single-copy element in the CLS strains, but duplicated in strain 9a5c). Moreover, there was one prophage (XfP6) common to strains 9a5c and 6c only. Finally, strain 6c bore a 9.5 kb phage remnant that was not present in any other strain (prophage XfP6c-1), whilst strain 32 bore two exclusive elements: a large, 35.1 kb, complete prophage (prophage XfP32-1) and a smaller, 34.9 kb element (prophage XfP32-2), which might be a phage remnant, related to XfP32-1 (Fig. 4). As usually observed, all these newly identified elements seemed to carry genes from other bacterial species, contributing to enlarging the flexible gene pool of their specific hosts, with possible impacts on their capacity to interact with different environments (Gal-Mor & Finlay, 2006). For example, XfP6c-1 carried a homologue of the zot (zonula ocludens toxin) gene – a gene that has already been described in other *X. fastidiosa* strains and that may play a role in disease development (Hagemann et al., 2006; da Silva et al., 2007); prophage XfP32-1 carried a luxR transcriptional regulator, which may affect gene expression control in the

**Table 2. ANI values derived from X. fastidiosa pairwise comparisons**

ANI values obtained after comparisons involving the genomes of seven *X. fastidiosa* strains obtained from both North America (M12, M23, Temecula1 and Ann-1) and South America (9a5c, 32 and 6c). Comparisons were also made with the genome of *Xanthomonas albilineans* (strain GPE PC73) to provide a reference ANI value with an unrelated bacterium.

<table>
<thead>
<tr>
<th>Strain</th>
<th>32</th>
<th>6c</th>
<th>M12</th>
<th>Temecula1</th>
<th>M23</th>
<th>Ann-1</th>
<th>GPE PC73</th>
</tr>
</thead>
<tbody>
<tr>
<td>9a5c (<em>X. fastidiosa subsp. pauca</em>)</td>
<td>99.17</td>
<td>98.9</td>
<td>96.07</td>
<td>95.86</td>
<td>95.85</td>
<td>95.82</td>
<td>70.19</td>
</tr>
<tr>
<td>32 (<em>X. fastidiosa subsp. pauca</em>)</td>
<td>–</td>
<td>98.83</td>
<td>95.94</td>
<td>95.89</td>
<td>95.9</td>
<td>95.68</td>
<td>69.94</td>
</tr>
<tr>
<td>6c (<em>X. fastidiosa subsp. pauca</em>)</td>
<td>–</td>
<td>96.16</td>
<td>95.88</td>
<td>95.85</td>
<td>95.91</td>
<td>69.94</td>
<td></td>
</tr>
<tr>
<td>M12 (<em>X. fastidiosa subsp. multiplex</em>)</td>
<td>–</td>
<td>97.55</td>
<td>97.53</td>
<td>97.46</td>
<td>69.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temecula1 (<em>X. fastidiosa subsp. fastidiosa</em>)</td>
<td>–</td>
<td>–</td>
<td>99.9</td>
<td>98.02</td>
<td>69.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M23 (<em>X. fastidiosa subsp. fastidiosa</em>)</td>
<td>–</td>
<td>–</td>
<td>98.0</td>
<td>–</td>
<td>69.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ann-1 (<em>X. fastidiosa subsp. sandyi</em>)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>69.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Xanthomonas albilineans</em> GPE PC73</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
bacterial host cell (van Kessel et al., 2013); and, finally, prophage $Xf^P32-2$ carried two new pilA genes, which encode the major components of type IV fimbriae – typically involved in adhesion and virulence in several pathogenic bacteria (Fronzes et al., 2008). The complete list of genes that could be found in these newly identified prophages, as identified through the NCBI Prokaryotic Genome Automatic Annotation Pipeline (see Methods), can be found in File S1 (available in the online Supplementary Material).

Interestingly, only two of the prophages shared by all South American X. fastidiosa representatives could be found at conserved chromosomal positions: prophages $Xf^P3$ and $Xf^P5$, which were always located at the beginning and end of Block E, respectively (Fig. 3). The other elements appeared to have been involved in transposition during the course of evolution of South American X. fastidiosa strains, as their relative chromosomal positions changed in all three isolates compared herein. For example, prophage $Xf^P4$ was present at two distinct positions within Block E, in the cases of strains 9a5c and 32, whilst it appeared at the end of Block C, in strain 6c (Fig. 3). A similar situation was observed with prophage $Xf^P1$, which was the only element present in the genomes of South American strains that could also be found in Temecula1, where it was named prophage $Xf^PD4$ (de Mello Varani et al., 2008). Prophage $Xf^P1$ mapped to a position in the middle of Block D in the genome of strain 6c, but at the beginning of Block C in both strains 9a5c and 32 (although in opposite orientations). Moreover, strain 9a5c also bore prophage $Xf^P2$, which appeared to be a duplication of prophage $Xf^P1$ (Figs 3 and 4). Finally, prophage $Xf^P6$, which could be found in both strains 9a5c and 6c, also mapped at a...
different chromosomal location when these two strains were compared (Fig. 3).

Plasmid-related elements, which have also been found to constitute the flexible gene pool of *X. fastidiosa*, also contributed to differentiating the coffee-infecting strains. As mentioned above, the microarray hybridization analyses suggested that the plasmid-derived insertion element PI1 was present only in the genomes of group I CLS strains – information that was confirmed by the sequencing of strain 32. Moreover, sequencing of strain 6c confirmed the presence of a large, 39 kb plasmid (named p*xf*6c) that was extremely similar to p*xf*51, the large plasmid found in strain 9a5c. Moreover, gel electrophoresis of whole DNA samples obtained from strain 6c and other group II CLS strains confirmed the presence of a large episomal element in their genomes (data not shown). A direct comparison between p*xf*51 and p*xf*6c (Fig. 3b) showed that the major difference between them involved an ~15 kb region, located in the middle of p*xf*51, that was absent from p*xf*6c. This region carried a series of hypothetical proteins, as well as an operon constituted of seven Trb conjugation-associated genes (File S2).

**Gene content of the CLS-associated *X. fastidiosa* strains**

As mentioned above, the draft genomes obtained for strains 6c and 32 had ~2.56 and ~2.6 Mbp, respectively, likely to correspond to >95 % of the full genomic sequences for these two bacterial strains, as the fully sequenced genomes of other *X. fastidiosa* strains resulted in chromosomes with sizes that ranged from 2.39 to 2.73 Mbp (Simpson et al., 2000; Bhattacharyya et al., 2002; Van Sluys et al., 2003; Chen et al., 2010; Schreiber et al., 2010; Zhang et al., 2011). Moreover, base quality of the resulting consensus sequences for both draft genomes exceeded Phred scores of 40 in >95 % of the bases (data not shown). Thus, these draft sequences provided sufficiently reliable information to allow identification of most ORFs present in the genomes of the two CLS-related *X. fastidiosa* strains.

Thus, the draft genomes of strains 6c and 32 were annotated through submission to the NCBI Prokaryotic Genome Automatic Annotation Pipeline (see Methods), resulting in preliminary ORF lists for each strain. Although completion of the genomes will certainly lead to an improvement in these numbers, the current annotations identified 2506 ORFs in the genome of strain 6c and 2477 ORFs in the genome of strain 32. Comparative analyses regarding gene content and similarity were made with the genomes of all four strains shown in Fig. 3 (using Mauve), resulting in a preliminary pangenome summarized in Fig. 5. A complete list of genes present in each *X. fastidiosa* strain (as well as its orthologues) can be found in File S2.

This analysis confirmed that the four strains shared a core genetic pool composed of ~1600 genes, responsible for the major metabolic routes and main structural features of the bacterial cell. Nonetheless, a large number of genes found in this pangenome were either present exclusively in one strain or shared by subgroups of *X. fastidiosa* strains. Moreover, there were many cases of genes that, although present in more than one strain, showed reduced similarity amongst themselves, meaning that they were likely to exhibit structural differences that may impact their respective biological functions. Although we realize that it is not possible to perform a complete comparison amongst members of this flexible gene pool with annotations made from draft genomes, preliminary analyses derived from data contained in Fig. 5 and File S2 provided interesting information regarding differences in genes that were implicated in host specificity and virulence in many pathogenic bacteria, including *X. fastidiosa*. The structural specificities found in some of these genes, as well as a more detailed evaluation of the biological consequences that may be derived from these and other variations in gene content, are further addressed in the Discussion.

**DISCUSSION**

Recent research, stimulated by economic interests, has turned *X. fastidiosa* into an intensively studied biological model, but many aspects regarding the evolutionary history of the group, as well as mechanisms that govern disease development and host specificity, are yet to be elucidated (Simpson et al., 2000; Bhattacharyya et al., 2002; Chen et al., 2010). Comparative genomic analyses have been performed with a great variety of strains of this phytopathogen, providing clues that are helping to unravel the basis of genomic evolution and adaptation of these bacteria to interact with such a broad range of hosts and induce the various diseases attributed to *X. fastidiosa* strains (Nunes et al., 2003; Doddapaneni et al., 2006, 2007; Varani et al., 2012). The present study brings to this puzzle, for the first time, genomic information regarding strains responsible for CLS – a pathogeny that has been known to occur in South America since 1997 (Paradelo-Filho et al., 1997).

The genomic information regarding CLS-associated *X. fastidiosa* strains confirms previous information concerning the importance of HTEs as major players in the evolution of this group of pathogens (Nunes et al., 2003; de Mello Varani et al., 2008). As mentioned earlier, previous studies involving *X. fastidiosa* strains obtained from other plant hosts isolated from diverse geographical origins confirmed that different bacterial strains carry a large amount of laterally transferred elements (mostly prophages) (Varani et al., 2013). The genomic information regarding the CLS-associated *X. fastidiosa* strains fits well into this scenario, as the two sequenced strains display ~8.64 % of sequences associated with such elements throughout their respective chromosomal sequences. Moreover, three new phage elements can be found in these bacteria.

Van Sluys et al. (2003) highlighted the importance of prophages as mediators of chromosomal rearrangements.
Fig. 4. Comparisons amongst corresponding prophage elements found in the genomes of different X. fastidiosa strains. (a) Alignment of prophages XfP1/XfP2 (described in strain 9a5c by Simpson et al., 2000) and their corresponding elements in strains 32 and 6c. (b) Alignment of prophage XfP3 (described in strain 9a5c by Simpson et al., 2000) and the corresponding elements in strains 32 and 6c. (c) Alignment of prophage XfP4 (described in strain 9a5c by Simpson et al., 2000) and the...
amongst *X. fastidiosa* strains, suggesting that the alignment patterns verified in the chromosomes of strains 9a5c and Temecula1 could be interconverted by two chromosomal inversions and one recombination event, probably caused by prophages. Interestingly, the chromosomal structures verified in CLS-related strains depict a syntenic pattern that appears to represent an intermediary structure between the Temecula1 and 9a5c chromosomes. As shown in Fig. 3, the basic chromosomal features of the coffee-infecting strains can be rebuilt from the Temecula1 chromosome by a reciprocal recombination (followed by inversion) of Blocks D and F + B, which could have been mediated by prophages XfP3 and XfP4, which appear at the beginning and end of Block D in the genome of strain 6c (see Fig. 3). Similarly, the syntenic arrangement of strain 9a5c can be explained by the simple inversion of a large chromosomal region (involving blocks B + C + D + E + F), which is likely to have occurred during the evolution of citrus-infecting bacteria from an ancestral coffee-related strain. This chromosomal inversion might have been mediated by insertion/recombination events involving prophages XfP1/ XfP2, as they are found at the ends of this inverted block in the genome of strain 9a5c. Interestingly, this possibility is in accord with the hypothesis that coffee-related bacteria originated from North American *X. fastidiosa* strains and originated the citrus-infecting bacteria (Almeida et al., 2008; Nunney et al., 2012). Additionally, prophage transposition may also have assisted in the evolution of *X. fastidiosa* strains as, as observed in Fig. 3, some of these elements appear to have changed their relative positions along the chromosomes of even closely related bacteria. This phenomenon is likely to have occurred as a consequence of chaperone-assisted excisive recombination – a mechanism of lysogeny escape verified in many temperate bacteriophages that affects not only fully functional prophages, but also phage remnants, which can transpose to other regions of the bacterial chromosome as a result of the interaction between prophage-encoded factors and host enzymes, such as the molecular chaperone DnaJ (Puvirajesinghe et al., 2012). This possibility is favoured by the observation that gene expression analysis of *X. fastidiosa* cells showed that growth conditions could lead to strong and coordinated upregulation of practically all genes present in prophage sequences of strain 9a5c, providing the cells with the necessary phage-specific factors that mediate excision (Nunes et al., 2003). Other experiments demonstrated that prophage elements present in the genomes of different *X. fastidiosa* strains could become transcriptionally active in response to external stimuli, including chemical conditions that mimic xylem sap (Ciraulo et al., 2010; Varani et al., 2013).

Comparisons involving differences in gene content also identified structural variations in genes that have been suggested to be important players in host recognition and disease development in many different bacteria (Fig. 6). For instance, *X. fastidiosa* has been shown to possess genes encoding both type I (*fim*) and type IV (*pil*) fimbriae, and these two structures have been confirmed as important virulence factors in this phytopathogen, as well as in other pathogenic bacteria (Meng et al., 2005; Weissman et al., 2006; Li et al., 2007; Fronzes et al., 2008; Cursino et al., 2011; Aagesen & Häse, 2012). Type I fimbriae (composed essentially of polymerized FimA subunits) play a central role in cell attachment, aggregation and biofilm formation (Blumer et al., 2005). Type IV pilus, however, are composed essentially of PilA subunits and have been shown to be involved in twitching motility – an important mechanism that helps non-flagellated bacteria to move along substrates (Nguyen et al., 2012).

Initial studies by Meng et al. (2005) demonstrated that both types of fimbriae were capable of mediating *X. fastidiosa*–host interactions, playing key roles in cell aggregation, substrate attachment, biofilm formation and tissue colonization. Further studies reported by Li et al. (2007) demonstrated that a fimA *X. fastidiosa* mutant (deficient in type I fimbriae) was not capable of producing biofilm, whilst mutations that affected the pil genes produced bacteria that were not capable of forming type IV fimbriae, displaying reduced or impaired twitching

---

**Fig. 5.** Venn diagram showing the distribution of ORFs exclusive/common to the genomes of the four *X. fastidiosa* strains compared in this work. ORF compositions of strains 32 and 6c were obtained from the whole-genome shotgun projects deposited at GenBank (see Methods). ORF compositions of strains 9a5c and Temecula1 were obtained from RefSeq IDs NC_002488.3 and NC_004556.1, respectively.
motility, which affected their capacity to colonize xylem vessels upstream the flow of sap from the point of inoculation. The importance of FimA and PilA fimbrial adhesins as determinants of infective capacity and biofilm formation in *X. fastidiosa* cells has also been demonstrated by microarray-assisted gene expression analyses (de Souza et al., 2003, 2004). As highlighted in Fig. 6(a), the four *X. fastidiosa* strains share two major fimA genes, in which small degrees of structural variability can be observed. However, the three South American strains bear a third locus, constituted by XF0080, 6c-00310 and 32-00310. Interestingly, although Temecula1 presents the gene corresponding to this third locus (PD0058), a frameshift mutation renders this gene non-functional. Moreover, another isomorph of this gene can be found in strains 9a5c and 6c (consisting of ORFs XF0077 and 6c-00300), indicating that significant structural variability may be observed amongst the type I fimbriae of different *X. fastidiosa* strains. The pilA genes have been shown to display a similar situation, as observed in Fig. 6(b): the four *X. fastidiosa* strains seem to share two major pilA loci: locus 1 is represented by XF2542, 6c-11285, 32-11515 and PD1926, whilst locus 2 involves XF2539, 6c-11275, 32-11505 and PD1924. However, additional pilA isomorphs can be found in seven additional chromosomal loci throughout these bacterial genomes, suggesting that the four *X. fastidiosa* strains analysed in this study may also carry structural variability in their type IV fimbriae. Altogether, it is tempting to speculate if the occurrence of structural variability in both types of fimbriae may correlate with distinct binding specificities, influencing *X. fastidiosa* host specificity and colonization, as observed in other pathogenic bacteria (Weissman et al., 2006; Aagesen & Häse, 2012).

**Fig. 6.** Similarity dendrograms showing the evolutionary relatedness and degree of sequence similarity amongst some paralogous groups of genes (shown to be involved in processes of host colonization and/or virulence in other pathogenic bacteria) that display structural differences amongst the four *X. fastidiosa* strains analysed in this study. Genes present in the same locus, in different strains, were identified by Mauve (Darling et al., 2010) and are displayed in similarly coloured branches. Predicted protein sequences were aligned with the aid of CLUSTAL W (Larkin et al., 2007) and the resulting dendrograms were reconstructed with the aid of PROTDEND, part of PHYLIP version 3.695 (Felsenstein, 1989). The resulting analyses are shown for all orthologues encoding (a) fimbrillin (*fimA*), (b) pilin (*pilA*) and (c) haemolysin genes. The scale bar indicates the number of sequence changes measured during the evolution of the protein sequences under comparison.
Another important difference regarding gene composition amongst X. fastidiosa strains involves elements responsible for the production of LPS, which account for ~75% of the bacterial outer membrane structure, and is known to mediate interactions between the bacterial cell and its surrounding environment (Greenfield & Whitfield, 2012; Kannenberg & Carlson, 2001; Thanweer et al., 2008). So far, preliminary characterization of X. fastidiosa LPS has been performed only for T. emella1 (Clifford et al., 2013), which demonstrated that a rhamnose-rich O-antigen is crucial to surface attachment, cell–cell aggregation and biofilm maturation. Moreover, depletion of the O-antigen moiety compromises the bacterium’s ability to colonize the host (Clifford et al., 2013). Previous comparative analyses demonstrated that the South American strain 9a5c contained a gene encoding an O-antigen acetylase (oafA), which is absent in all North American strains sequenced so far (Bhattacharyya et al., 2002; Van Sluys et al., 2003; Chen et al., 2010). Analysis of the CLS-related X. fastidiosa genomes shows that these strains also carry a copy of oafA, confirming that differences in LPS are likely to bear significant structural differences between South and North American X. fastidiosa strains (Thanweer et al., 2008). OafA is homologous to a series of proteins that define a family of membrane factors involved in the acylation of carbohydrate moieties on extracellular molecules, including the NodX from Rhizobium leguminosarum, which is responsible for acetylation of the Nod factors, which determine host range specificity in this plant-colonizing bacteria (Kannenberg & Carlson, 2001; Lerouge et al., 2003).

However, the genomic comparisons presented in this study indicate that LPS structural variation may be even more complex amongst South American X. fastidiosa strains as strain 6c bears a frameshift mutation that inactivates the gene that encodes a UDP-glucurionate decarboxylase, which is involved in the synthesis of UDP-xylose— the sugar-nucleotide precursor that donates xylose moieties to be incorporated into the O-antigen side chain of bacterial LPS (Gu et al., 2010) (File S3). At this point, we do not know if this mutation is fortuitously present only in strain 6c or if it represents a genetic condition widespread throughout the entire group II of coffee-related strains. If that is the case, it will be interesting to verify if the differences in their LPS structures may act as determinants for host specificity and/or virulence amongst these bacteria.

Divergence has also been observed amongst genes that encode calcium-binding haemolysin-like proteins, which have been identified in all four X. fastidiosa strains (Fig. 6c). Haemolysins belong to a family of related exotoxins with cytolytic activity known as RTX toxins, whose members can be found in many different bacteria—including phytopathogens, where they have been demonstrated to enhance bacterial competitiveness against other endophytes (Oresnik et al., 1999). A final difference in gene content that deserves to be highlighted, due to its potential involvement in the process of host colonization, is the structure of the polygalacturonase gene. Previous studies have shown that a frameshift mutation, found in strain 9a5c, renders polygalacturonase inactive in this particularly strain, and this deficiency has always been speculated to correlate with the different infective capacities of strains 9a5c and T. emella1 as this gene may be associated with degradation of the pit membranes that separate elements of the plant xylem (Simpson et al., 2000; Van Sluys et al., 2003; Sun et al., 2011; Roper et al., 2007). Characterization of the CLS-related strains shows that the same frameshift mutation inactivates the polygalacturonase gene in strain 32, but not in strain 6c, which bears a fully functional copy of this gene (File S4).

Finally, the draft genomes of strains 6c and 32 allowed us to obtain more precise information regarding variations in the ANI values amongst X. fastidiosa strains. ANI values have been gaining increasing credibility over classic DNA hybridization experiments as a tool to assist in the taxonomic identification of bacterial species (Goris et al., 2007). The current paper shows that all pairs of North American X. fastidiosa strains displayed ANI numbers that range between 97.46 and 99.9%, indicating a very close phylogenetic proximity and supporting the classification of all these bacteria as a single species. In a similar way, all pairs of South American X. fastidiosa displayed ANI values >98%. However, ANI calculations involving any pair of South/North American strains displayed ANI values that are very close to (mostly below) 96%, which is currently accepted to delimit the threshold between species (Konstantinidis & Tiedje, 2005; Goris et al., 2007; Richter & Rossello-Móra, 2009). Thus, South American X. fastidiosa strains, currently classified as members of X. fastidiosa subsp. pauca, whilst closely related amongst themselves, display a significantly distant phylogenetic relationship with their North American counterparts, raising questions regarding the positioning of all X. fastidiosa strains as members of the same species.

In summary, here we provide, for the first time, a comprehensive genomic analysis of coffee-infecting X. fastidiosa strains. Data have been obtained from representatives belonging to two phylogenetically distinct groups of CLS-related bacteria and confirm the importance of HTEs as major mediators of chromosomal evolution amongst these bacteria. Moreover, significant differences regarding gene content were identified when these bacteria were compared with previously sequenced X. fastidiosa strains, obtained from alternative plant hosts. Although direct experimentation still needs to be performed to elucidate their biological consequences, it was of interest to verify that CLS-related bacteria display variations in genes that produce toxins (which may help them to compete with endophytes present in their specific hosts) as well as surface-related factors (such as fimbrial adhesins and LPS) that have been shown to be involved with recognition of specific host factors in different pathogenic bacteria, influencing their host specificity, infectivity and/or development of virulence capacity.
ACKNOWLEDGEMENTS

The authors wish to thank Dr Joao R.S. Lopes (Department of Entomology and AcaRCol, ESALQ/Universidade de São Paulo) for providing some of the X. fastidiosa strains used in this study. This work was supported by grants from Fundação de Amparo à Pesquisa do Estado de São Paulo (09/50114-7 and 09/50115-3) and Fundação de Amparo ao Ensino e à Pesquisa. D.S.S. received a scholarship grant from Fundação de Amparo à Pesquisa do Estado de São Paulo (08/53231-1). D.B., V.C.A. and A.C.F.O. received scholarship grants from the Brazilian Federal Agency CAPES.

REFERENCES


Comparative genomic analysis of X. fastidiosa

Pieretti, I., Royer, M., Barbe, V., Carrere, S., Koebnik, R., Cociancich, S., Couloux, A., Darras, E., Gouzy, J. & other authors (2009). The complete genome sequence of Xanthomonas albilinea provides new insights into the reductive genome evolution of the xylem-limited Xanthomonadaceae. BMC Genomics 10, 616.


Edited by: G. Preston