High-Quality Draft Genome Sequence Resources of Eight *Xylella fastidiosa* Strains Isolated from Citrus, Coffee, Plum, and Hibiscus in South America

Paulo Marques Pierry,¹ Wesley Oliveira de Santana,¹ João Paulo Kitajima,² Joaquim Martins-Junior,¹ Paulo Adriano Zaini,^{1,3} Guillermo Uceda-Campos,¹ Oseias R. Feitosa-Junior,¹ Patrícia Isabela Silva Pessoa,¹ Helvécio Della Coletta-Filho,⁴ Alessandra Alves de Souza,⁴ Marcos Antonio Machado,⁴ Abelmon da Silva Gesteira,⁵ Layla Farage Martins,¹ Murilo Sena Amaral,¹ Felipe Cesar Beckedorff,¹ Luiz Gonzaga Paula de Almeida,⁶ Ana Tereza Ribeiro de Vasconcelos,⁶ Sergio Verjovski-Almeida,¹ João Carlos Setubal,¹ and Aline Maria da Silva^{1,†}

¹ Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, São Paulo, SP, Brazil

² Mendelics Análise Genômica, São Paulo, SP, Brazil

³ Department of Plant Sciences, University of California, Davis, CA, U.S.A.

⁴ Centro de Citricultura Sylvio Moreira, Instituto Agrônomo de Campinas, Cordeirópolis, SP, Brazil

⁵ Embrapa Mandioca e Fruticultura, Cruz Das Almas, BA, Brazil

⁶ Laboratório Nacional de Computação Científica, Petrópolis, RJ, Brazil

Abstract

Xylella fastidiosa subsp. *pauca*, once confined to South America and infecting mainly citrus and coffee plants, has been found to be associated with other hosts and in other geographic regions. We present high-quality draft genome sequences of *X. fastidiosa* subsp. *pauca* strains J1a12, B111, U24D, and XRB isolated from citrus plants in Brazil, strain Fb7 isolated from a citrus plant in Argentina and strains 3124, Pr8x, and Hib4 isolated, respectively, from coffee, plum, and hibiscus plants in Brazil. Sequencing was performed using Roche 454-GS FLX, MiSeq-Illumina or Pacific Biosciences platforms. These high-quality genome assemblies will be useful for further studies about the genomic diversity, evolution, and biology of *X. fastidiosa*.

Genome Announcement

Xylella fastidiosa is a xylem-limited Gram-negative bacterium (Wells et al. 1987) that is transmitted among plant hosts by xylem sap-feeding insects (Redak et al. 2004). Strains of this bacterium are the causal agent of perennial crop diseases such as Pierce's disease of grapevine (Hopkins 1989), citrus variegated chlorosis (CVC) (Rossetti et al. 1990), coffee leaf scorch (CLS) (de Lima et al. 1998), plum leaf scald (PLS) (Raju et al. 1982), and olive quick decline syndrome (OQDS) (Saponari et al. 2017). Although *X. fastidiosa* has been associated

Funding

Funding for this work was provided by São Paulo Research Foundation (FAPESP), research grants 08/11703-4, 09/13527-1, 11/01217-8, and 14/50880-0, by Coordination for the Improvement of Higher Education Personnel (CAPES), research grant 3385/2013, and by National Council for Scientific and Technological Development (CNPq), research grant 465440/2014-2. P. M. Pierry and J. Martins-Junior received fellowships from FAPESP (grants 09/13527-1 and 11/01217-8). W. O. de Santana, G. Uceda-Campos, and O. R. Feitosa-Junior were supported by fellowships from CAPES. H. D. Coletta-Filho, A. A. de Souza, M. A. Machado, A. S. Gesteira, A. T. R. de Vasconcelos, S. Verjovski-Almeida, J. C. Setubal, and A. M. da Silva received Research Fellowship Awards from CNPq.

Keywords

bacteriology, citrus variegated chlorosis, phytopathogen, *Xylella fastidiosa* subsp. *pauca*

[†]Corresponding author: A. M. da Silva; almsilva@iq.usp.br

Current address of W. Oliveira de Santana: Universidade Estadual do Piauí, Picos, PI, Brazil.

Current address of J. Martins-Junior: Centro Nacional de Pesquisa em Energia e Materiais, Campinas, SP, Brazil.

Current address of P. I. S. Pessoa: Instituto Federal de São Paulo, São Roque, SP, Brazil.

Current address of M. S. Amaral and S. Verjovski-Almeida: Laboratório de Parasitologia, Instituto Butantan, São Paulo, SP, Brazil.

Current address F. C. Beckedorff: Department of Human Genetics, Sylvester Comprehensive Cancer Center, University of Miami Miller School of Medicine, Miami, FL, U.S.A.

The author(s) declare no conflict of interest.

Accepted for publication 8 June 2020.

Table 1. Assembly metrics and selected features of eight *Xylella fastidiosa* sequenced genomes^a

Strain	Reference	Host of origin	Location	Sequence type ^b	Sequenced bases (Mbp)			Average genome coverage ^c
					Shotgun	454 Roche paired end tags	PacBio	
J1a12	Monteiro et al. (2001)	<i>Citrus sinensis</i> 'Pêra'	São Paulo (Brazil)	ST11	102 ^d	74.7	NA	65.9x
B111	Monteiro et al. (2001)	<i>Citrus sinensis</i> 'Pêra'	São Paulo (Brazil)	ST11	387 ^e	NA	NA	144.4x
U24D	This work	<i>Citrus sinensis</i> 'Baianinha'	São Paulo (Brazil)	ST13	200 ^d	19.1	NA	81.7x
Fb7	da Silva et al. (2007)	<i>Citrus sinensis</i> 'Valencia'	Corrientes (Argentina)	ST69	NA	60.6	1174	460x
XRB	This work	<i>Citrus sinensis</i> 'Pêra'	Bahia (Brazil)	ST11	429 ^e	NA	NA	160x
3124	Monteiro et al. (2001)	<i>Coffea arabica</i>	São Paulo (Brazil)	ST16	723 ^d	NA	NA	269.8x
Pr8x	This work	<i>Prunus</i>	São Paulo (Brazil)	ST14	171 ^d	NA	NA	63.8x
Hib4	This work	<i>Hibiscus</i>	Distrito Federal (Brazil)	ST70	268 ^d	NA	NA	100x

(Continued on next page)

^a NA indicates not available/not applicable.^b Scally et al. (2005).^c 454/Roche shotgun reads.^d Illumina/MiSeq shotgun paired-end reads.^e Average genome coverage was estimated using the genome of *X. fastidiosa* reference strain 9a5c (2,679,305 bp).^f Estimated size.

with disorders in several other plant species (Hopkins and Purcell 2002), most of its host plant species are asymptomatic (Sicard et al. 2018).

Strains of *X. fastidiosa* are currently categorized into five subspecies: *fastidiosa*, *pauca*, *multiplex*, *sandyi*, and *morus*, which are presumed to have originated in Central America (subsp. *fastidiosa*), South America (subsp. *pauca*) and North America (subsp. *multiplex*, *sandyi*, and *morus*) (Almeida and Nunney 2015; Sicard et al. 2018). Once confined to South America and infecting mainly citrus and coffee plants, *X. fastidiosa* subsp. *pauca* has been found associated with other hosts such as olive plants and in other geographic regions (e.g., Southern Italy, Mallorca Island, and Costa Rica) (Nunney et al. 2014; Safady et al. 2019; Saponari et al. 2017).

In this work, we present high-quality draft genome sequences of eight strains of *X. fastidiosa* subsp. *pauca* isolated from symptomatic plants in South America. The strains J1a12, B111, 3124 (Monteiro et al. 2001), U24D, XRB, Pr8x, and Hib4 were isolated in Brazil, while strain Fb7 (da Silva et al. 2007) was isolated in Argentina (see Table 1 for their host of origin and location of isolation). No information regarding the pathogenic phenotype of strains U24D, XRB, Fb7, 3124, Pr8x, and Hib4 is available. On the other hand, it has been reported that strain B111 can multiply and produce typical CVC symptoms in citrus plants (Teixeira et al. 2004), while strain J1a12 was shown to be nonpathogenic to *Citrus sinensis* and *Nicotiana tabacum* (Koide et al. 2004). Both J1a12 and B111 are suitable to genetic transformation (Monteiro et al. 2001; Teixeira et al. 2004).

Whole-genome sequencing of each strain was performed using Roche 454-GS FLX, MiSeq-Illumina, or Pacific Biosciences (PacBio, Menlo Park, CA) platforms (Table 1). *X. fastidiosa* strains were cultured in 50 ml of PW broth (Davis et al. 1981) plus 0.5% glucose for 7 days at 28°C in a rotary shaker, until OD_{600nm} of 0.8 to 1.2. Cells were harvested (12,000 × g for 5 min) and total DNA was extracted using Wizard Genomic DNA Purification Kit (Promega Corporation, Madison, WI) followed by a clean-up step using QIAamp mini spin columns (Qiagen, Hilden, Germany). DNA purity and concentration were evaluated by the absorbance at 260, 280, and 230 nm on a NanoDrop ND-2000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA). DNA concentration was further quantified with Quant-iT Picogreen dsDNA assay kit (Thermo Fisher Scientific). Purified DNA samples were enriched in fragments larger

Table 1. (Continued from previous page)

Completeness (%)	Contamination (%)	Number of contigs	N50 (bp)	%GC (whole genome)	Chromosome size (bp)	Plasmid size (bp)	NCBI accession number	IMG/M genome ID
99.59	0.18	3	NA	52.81	2,788,789	51,180	NZ_CP009823.1; NZ_CP009825.1	2510065065; 2510065067
99.64	0.18	77	130,109	52.36	2,682,276 ^f	27,268 50,444	NZ_CP009824.1 GCF_013283685	2510065066 2565956513; contig 9_plasmid.76 contig 5_plasmid.77
99.59	0.18	2	NA	52.68	2,681,334	51,156	NZ_CP009790.1; NZ_CP009791.1	2513237126; 2513237131
98.78	0.00	2	NA	52.51	2,659,912	39,408	NZ_CP010051.2; NZ_CP014330.2	2695421029
99.59	0.18	74	138,935	52.44	2,714,643 ^f	48,841	GCF_013283695	2565956512; contig 29.74 (plasmid) contig 5.73 (plasmid)
99.64	0.00	1	NA	52.63	2,748,594	NA	NZ_CP009829.1	2513237125
99.59	0.18	2	NA	52.59	2,666,242	39,580	NZ_CP009826.1; NZ_CP009827.1	2513237127; 2513237130
99.64	1.45	2	NA	52.59	2,813,297	64,251	NZ_CP009885.1; NZ_CP009886.1	2513237128; 2513237129

than 10 kbp DNA according to electrophoresis on a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA).

DNA samples (~5 µg) of strains J1a12, U24D, 3124, Pr8x, and Hib4 were nebulized to obtain fragments ranging from 500 to 800 bp and shotgun libraries were constructed using Roche 454 GS Titanium Rapid Library Prep Kit DNA. The 454-pyrosequencing was performed according to Roche 454 GS FLX Titanium standard protocols (Roche Applied Science, Penzberg, Germany). Strains J1a12, U24D and FB7 were also sequenced using Paired-End Tags strategy (Fullwood et al. 2009) with the same Roche 454 GS platform. For that, DNA samples (~15 µg) were sheared using the Hydroshear equipment (Digilab, Hopkinton, MA) to generate ~3 kbp fragments that were circularized prior to shotgun library construction. Strain Fb7 genome was resequenced using the PacBio RS II sequencing platform (PacBio). Briefly, ~15 µg of DNA was used to prepare a single 15 to 20 kbp library following the Pacific Biosciences library preparation protocol, loaded on a single-molecule real-time cell, and sequenced with C4 chemistry on PacBio RS II instrument. Shotgun libraries of strains B111 and XRB were prepared with 35 ng of DNA using the Illumina Nextera DNA library preparation kit (Illumina, San Diego, CA). The libraries were paired-end sequenced with MiSeq-Illumina platform using MiSeq Reagent kit v2 (500-cycle format).

De novo genome assemblies were performed with various computational tools. For strains J1a12, U24D, and Fb7, the reads resulting from the different sequencing platforms were combined for final assembly. Roche 454 sequencing reads were quality-filtered (quality score ≥30) and assembled using Newbler v.2.3 (454 Life Sciences). Assembly refinement was performed using MIRA 4 (Chevreux et al. 2004), Cross_Match (<http://www.phrap.org/phredphrap/general.html>), and manual interventions to solve genomic sequence breaks due to repetitive sequences. PacBio sequence reads were filtered and assembled with the PacBio Hierarchical Genome Assembly Process version 3 and polished using Miniasm (<https://github.com/lh3/miniasm>). Illumina paired-end reads (2 × 250 bp) were quality-filtered (quality score ≥30) and assembled using CLC Genomics Workbench (Qiagen, Hilden, Germany) using de novo assembly with further assembly refinements with the Microbial Genome Finishing Module.

The final assemblies for strains J1a12, U24D, Fb7, 3124, Pr8x, and Hib4 resulted in complete chromosomal sequences and their associated plasmid sequences (Table 1). Assemblies for strains B111 and XRB resulted in 77 and 74 contigs, respectively, encompassing the chromosome and two plasmids for each strain (Table 1). CheckM analysis (Parks et al. 2015) of these assemblies showed high completeness (>99.5%) and low contamination (<1.5%), indicating these are high-quality draft genomes sequences (Bowers et al. 2017). The average size of the eight genomes is 2.72 ± 0.06 Mbp (standard deviation) with the genome of strain Hib4 (2,813,297 bp) being the largest (Table 1).

Average nucleotide identities (ANI) of pairwise comparisons between the eight strains final chromosome assemblies were calculated using the Python module Pyani v0.2.7 (<https://github.com/widdowquinn/pyani>) (Pritchard et al. 2016) and were within the range of 98.47 to 99.93%. All eight genomes were submitted to the NCBI Prokaryotic Genome Annotation Pipeline (Tatusova et al. 2016) as well as to the IMG/M system (Chen et al. 2019) for automatic annotation. The complete chromosomes showed two copies of the 23S, 16S, and 5S rRNAs genes organized in two operons and 48 to 50 tRNA genes covering all the 20 amino acids. The number of protein coding genes for the eight strains ranged from 2,606 to 2,880 depending on the strain and annotation tool. The plasmids assembled in seven of these sequenced strains have been characterized in silico and empirically demonstrated (Pierry et al. 2020).

In silico analysis of the seven housekeeping genes (Sally et al. 2005) employed in *X. fastidiosa* multilocus sequence typing (<https://pubmlst.org/xfastidiosa/>) showed the sequenced strains belong to six different sequence types (ST): ST11 (J1a12, B111, and XRB), ST13 (U24D), ST14 (Pr8x), ST16 (3124), ST69 (Fb7), and ST70 (Hib4). The data presented here contribute to expand *X. fastidiosa* genome sequence data available and will be useful for further studies about the genomic diversity, evolution, and biology of this pathogen.

Availability of sequence data. All *X. fastidiosa* genome sequences described in this work have been deposited in GenBank/NCBI and IMG/M databases under the accession numbers listed in Table 1.

Acknowledgments

We thank Diva do Carmo Teixeira (*in memoriam*) from Fundo de Defesa da Citricultura for providing the strains B111 and J1a12. The plant material used for Hib4 strain isolation was provided by Elliott W. Kitajima (Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo). We also thank Rodrigo P. P. Almeida (Department of Environmental Science, Policy and Management, University of California, Berkeley) for fruitful suggestions.

Literature Cited

- Almeida, R. P. P., and Nunney, L. 2015. How do plant diseases caused by *Xylella fastidiosa* emerge? *Plant Dis.* 99:1457-1467.
- Bowers, R. M., Kyrpides, N. C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T. B. K., Schulz, F., Jarett, J., Rivers, A. R., Elae-Fadros, E. A., Tringe, S. G., Ivanova, N. N., Copeland, A., Clum, A., Becraft, E. D., Malmstrom, R. R., Birren, B., Podar, M., Bork, P., Weinstock, G. M., Garrity, G. M., Dodsworth, J. A., Yooseph, S., Sutton, G., Glockner, F. O., Gilbert, J. A., Nelson, W. C., Hallam, S. J., Jungbluth, S. P., Ettema, T. J. G., Tighe, S., Konstantinidis, K. T., Liu, W. T., Baker, B. J., Rattei, T., Eisen, J. A., Hedlund, B., McMahon, K. D., Fierer, N., Knight, R., Finn, R., Cochrane, G., Karsch-Mizrachi, I., Tyson, G. W., Rinke, C., Genome Standards, C., Lapidus, A., Meyer, F., Yilmaz, P., Parks, D. H., Eren, A. M., Schriml, L., Banfield, J. F., Hugenholtz, P., and Woike, T. 2017. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat. Biotechnol.* 35:725-731.
- Chen, I. A., Chu, K., Palaniappan, K., Pillay, M., Ratner, A., Huang, J., Huntemann, M., Varghese, N., White, J. R., Seshadri, R., Smirnova, T., Kirton, E., Jungbluth, S. P., Woike, T., Elae-Fadros, E. A., Ivanova, N. N., and Kyrpides, N. C. 2019. IMG/M v.5.0: An integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Res.* 47:D666-D677.
- Chevreaux, B., Pfisterer, B., Drescher, B., Driesel, A. J., Muller, W. E., Wetter, T., and Suhai, S. 2004. Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. *Genome Res.* 14: 1147-1159.
- da Silva, V. S., Shida, C. S., Rodrigues, F. B., Ribeiro, D. C., de Souza, A. A., Coletta-Filho, H. D., Machado, M. A., Nunes, L. R., and de Oliveira, R. C. 2007. Comparative genomic characterization of citrus-associated *Xylella fastidiosa* strains. *BMC Genomics* 8:474.
- Davis, M. J., French, W. J., and Schaad, N. W. 1981. Axenic culture of the bacteria associated with phony disease of peach and plum leaf scald. *Curr. Microbiol.* 6: 309-314.
- de Lima, J. E. O., Miranda, V. S., Hartung, J. S., Bransky, R. H., Coutinho, A., Roberto, S. R., and Carlos, E. F. 1998. Coffee leaf scorch bacterium: Axenic culture, pathogenicity, and comparison with *Xylella fastidiosa* of citrus. *Plant Dis.* 82:94-97.
- Fullwood, M. J., Wei, C. L., Liu, E. T., and Ruan, Y. 2009. Next-generation DNA sequencing of paired-end tags (PET) for transcriptome and genome analyses. *Genome Res.* 19:521-532.
- Hopkins, D. L. 1989. *Xylella fastidiosa*—Xylem-limited bacterial pathogen of plants. *Annu. Rev. Phytopathol.* 27:271-290.
- Hopkins, D. L., and Purcell, A. H. 2002. *Xylella fastidiosa*: Cause of Pierce's disease of grapevine and other emergent diseases. *Plant Dis.* 86:1056-1066.
- Koide, T., Zaini, P. A., Moreira, L. M., Vencio, R. Z., Matsukuma, A. Y., Durham, A. M., Teixeira, D. C., El-Dorry, H., Monteiro, P. B., da Silva, A. C., Verjovski-Almeida, S., da Silva, A. M., and Gomes, S. L. 2004. DNA microarray-based genome comparison of a pathogenic and a nonpathogenic strain of *Xylella fastidiosa* delineates genes important for bacterial virulence. *J. Bacteriol.* 186:5442-5449.
- Monteiro, P. B., Teixeira, D. C., Palma, R. R., Garnier, M., Bove, J. M., and Renaudin, J. 2001. Stable transformation of the *Xylella fastidiosa* citrus variegated chlorosis strain with oriC plasmids. *Appl. Environ. Microbiol.* 67:2263-2269.
- Nunney, L., Ortiz, B., Russell, S. A., Ruiz Sanchez, R., and Stouthamer, R. 2014. The complex biogeography of the plant pathogen *Xylella fastidiosa*: Genetic evidence of introductions and subspecific introgression in Central America. *PLoS One* 9:e112463.
- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. 2015. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25:1043-1055.
- Pierry, P. M., Uceda-Campos, G., Feitosa-Junior, O. R., Martins-Junior, J., de Santana, W. O., Della Coletta-Filho, H., Zaini, P. A., and da Silva, A. M. 2020. Genetic diversity of *Xylella fastidiosa* plasmids assessed by comparative genomics. *Trop. Plant Pathol.* 45:342-360.
- Pritchard, L., Glover, R. H., Humphris, S., Elphinstone, J. G., and Toth, I. K. 2016. Genomics and taxonomy in diagnostics for food security: Soft-rotting enterobacterial plant pathogens. *Anal. Methods* 8:12-24.
- Raju, B. C., Wells, J. M., Nyland, G., Bransky, R. H., and Lowe, S. K. 1982. Plum leaf scald—isolation, culture, and pathogenicity of the causal agent. *Phytopathology* 72: 1460-1466.
- Redak, R. A., Purcell, A. H., Lopes, J. R., Blua, M. J., Mizell, R. F., and Andersen, P. C. 2004. The biology of xylem fluid-feeding insect vectors of *Xylella fastidiosa* and their relation to disease epidemiology. *Annu. Rev. Entomol.* 49:243-270.

- Rossetti, V., Garnier, M., Bove, J. M., Beretta, M. J. G., Teixeira, A. R. R., Quaggio, J. A., and Denegri, J. D. 1990. Occurrence of xylem-restricted bacteria in sweet orange trees affected by chlorotic variegation, a new citrus disease in Brazil. *Comp. Rendus Acad. Sci.* 310:345-349.
- Safady, N. G., Lopes, J. R. S., Francisco, C. S., and Della Coletta, H. 2019. Distribution and genetic diversity of *Xylella fastidiosa* subsp. *pauca* associated with olive quick syndrome symptoms in southeastern Brazil. *Phytopathology* 109: 257-264.
- Saponari, M., Boscia, D., Altamura, G., Loconsole, G., Zicca, S., D'Attoma, G., Morelli, M., Palmisano, F., Saponari, A., Tavano, D., Savino, V. N., Dongiovanni, C., and Martelli, G. P. 2017. Isolation and pathogenicity of *Xylella fastidiosa* associated to the olive quick decline syndrome in southern Italy. *Sci. Rep.* 7: 17723.
- Scally, M., Schuenzel, E. L., Stouthamer, R., and Nunney, L. 2005. Multilocus sequence type system for the plant pathogen *Xylella fastidiosa* and relative contributions of recombination and point mutation to clonal diversity. *Appl. Environ. Microbiol.* 71:8491-8499.
- Sicard, A., Zeilinger, A. R., Vanhove, M., Scharrel, T. E., Beal, D. J., Daugherty, M. P., and Almeida, R. P. P. 2018. *Xylella fastidiosa*: Insights into an emerging plant pathogen. *Annu. Rev. Phytopathol.* 56:181-202.
- Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E. P., Zaslavsky, L., Lomsadze, A., Pruitt, K. D., Borodovsky, M., and Ostell, J. 2016. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* 44:6614-6624.
- Teixeira, D. C., Rocha, S. R., de Santos, M. A., Mariano, A. G., Li, W. B., and Monteiro, P. B. 2004. A suitable *Xylella fastidiosa* CVC strain for post-genome studies. *Curr. Microbiol.* 49:396-399.
- Wells, J. M., Raju, B. C., Hung, H. Y., Weisburg, W. G., Mandelcopaul, L., and Brenner, D. J. 1987. *Xylella fastidiosa* gen. nov., sp. nov.: Gram-negative, xylem-limited, fastidious plant bacteria related to *Xanthomonas* spp. *Int. J. Syst. Bacteriol.* 37:136-143.