UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE QUÍMICA
Programa de Pós-Graduação em Bioquímica

PAULO MARQUES PIERRY

**Pirossequenciamento e análise comparativa de genomas
do fitopatógeno *Xylella fastidiosa***

Versão corrigida da Dissertação defendida

São Paulo

Data de Depósito na SPG: 02/03/2012

PAULO MARQUES PIERRY

**Pirossequenciamento e análise comparativa de genomas do
fitopatógeno *Xylella fastidiosa***

Dissertação apresentada ao Instituto de
Química da Universidade de São Paulo para obtenção
do Título de Mestre em Ciências, área Bioquímica

Orientadora: Profa. Dra. Aline Maria da Silva

São Paulo
2012

**AUTORIZO A REPRODUÇÃO E A DIVULGAÇÃO DE TODO OU PARTE DESTES
TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA
FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.**

DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO DO CONJUNTO DAS QUÍMICAS

Pierry, Paulo Marques

Pirossequenciamento e análise comparativa de genomas do fitopatógeno
Xylella fastidiosa. / Paulo Marques Pierry. --São Paulo, 2012.

Orientadora: Aline Maria da Silva.

144 f.

Dissertação (Mestrado) – Universidade de São Paulo. Instituto de Química,
Programa de Pós-Graduação em Ciências Biológicas. Área de Concentração:
Bioquímica.

Descritores: 1. Pirossequenciamento. 2. Análise de genomas. 3. Escaldadura
da folha de ameixeira 4. Clorose Variegada de Citrus. 5. Escaldadura de
folha de *Hibiscus*. 6. *Xylella fastidiosa*. I. da Silva, Aline Maria II.
Universidade de São Paulo. Instituto de Química. Programa de Pós-
Graduação em Ciências Biológicas (Bioquímica). III. Título.

Paulo Marques Pierry

**Pirossequenciamento e análise comparativa de genomas do fitopatógeno
*Xylella fastidiosa***

Dissertação apresentada ao Instituto de
Química da Universidade de São Paulo
para obtenção do Título de Mestre em
Ciências, área Bioquímica.

Aprovado em: _____

Banca Examinadora

Prof. Dr. _____

Instituição: _____

Assinatura: _____

Prof. Dr. _____

Instituição: _____

Assinatura: _____

Prof. Dr. _____

Instituição: _____

Assinatura: _____

Dedico...

Aos meus pais e irmão, minha namorada e a todos que me ajudaram.

Agradeço...

À Profa. Dra. Aline Maria da Silva por aceitar me orientar e por todo o suporte e orientação que me deu durante todo o período de meu mestrado.

Aos Profs. Drs. Sergio Verjovski-Almeida, Walter Terra, Clélia Ferreira, Suely Lopes Gomes, Regina Baldini, Henning Ulrich, Glaucia Souza e Mauricio Baptista por possibilitarem acesso a equipamentos e reagentes para que eu pudesse realizar determinados experimentos.

A todos os professores que pude ter contato em disciplinas e seminários, dentro e fora do Instituto de Química.

Aos Drs. Marcos Machado, Helvécio Della-Coletta e Alessandra Alves, do Centro de Citricultura Sylvio Moreira, em Cordeirópolis, onde fui muito bem recebido e pude realizar parte do trabalho.

Aos funcionários do Centro de Citricultura Sylvio Moreira, em Cordeirópolis, que muito me ajudaram durante minha estadia.

À Profa. Dra. Sylvia Leão, da Universidade Federal de São Paulo, e sua aluna Michelle, as quais me receberam e me ajudaram durante experimentos lá realizados.

À Dra. Layla Farage, técnica do Centro Avançado de Tecnologias em Genômica (CATG) e colega de laboratório, pelo excelente trabalho no sequenciamento dos genomas de *Xylella*.

A Dra. Ana Tereza Vasconcelos e ao pesquisador Luis Gonzaga, do Laboratório Nacional de Computação Científica (Petrópolis, RJ) pela realização do sequenciamento utilizando a estratégia *Paired End Tags*, e pelo apoio em algumas análises de bioinformática.

Ao Dr. João Kitajima pela imensa atenção e dedicação, além de excelente trabalho na montagem dos genomas das cepas descritas neste trabalho.

Aos técnicos dos laboratórios em que frequentei e principalmente de meu laboratório, Alexandre Sanchez, grande amigo, Ivone e Dóris, por todo serviço prestado.

Aos colegas de laboratório Oséias, Renato, Patrícia, Rodrigo, Gustavo, Paulo Zaini, Luciana, Ana Paula, Joaquim, Fernando e Deibs pela amizade, ajuda em alguns experimentos e companhia durante o Mestrado.

Ao colega Wesley pela parceria neste trabalho, pela companhia nas viagens para Cordeirópolis e pela ajuda em vários experimentos.

Aos colegas de pós-graduação de outros laboratórios pela troca de experiências e apoio mútuo em horas difíceis.

Aos funcionários do IQUSP por proporcionarem um ambiente agradável de trabalho e aos serviços prestados.

Aos meus amigos de fora da USP, por todo o apoio e interesse durante o período de meu mestrado.

À minha namorada, Erika, pela maravilhosa companhia e imensurável apoio, carinho e amor em todos os momentos dessa minha trajetória.

Aos meus pais e família por me proporcionarem todo o suporte para que eu pudesse me dedicar inteiramente ao mestrado e por todo amor e apoio que me dedicam.

Ao CNPq, FINEP e FAPESP pelo apoio financeiro e à FAPESP pela bolsa de Mestrado.

RESUMO

Pierry, Paulo Marques. **Pirossequenciamento e análise comparativa de genomas do fitopatógeno *Xylella fastidiosa***. 2012. 144p. Dissertação de Mestrado - Programa de Pós-Graduação Ciências, área Bioquímica. Instituto de Química, Universidade de São Paulo, São Paulo.

Xylella fastidiosa é uma bactéria Gram-negativa, do subgrupo das Gama-Proteobactérias, não-flagelada, que coloniza o xilema de diversas plantas cultivadas e silvestres, podendo ser causadora de doenças. Sua disseminação é feita por insetos conhecidos como cigarrinhas. Genomas de cepas de *X. fastidiosa* isoladas de distintos hospedeiros já foram sequenciados completa ou parcialmente: 9a5c de citros; Temecula-1 e GB514 de videira; Dixon, M12 e M23 de amendoeira; Ann-1 de espiroleira e EB92-1, isolada de sabugueiro e utilizada como bio-controle para Doença de Pierce de videiras. Estudos de genômica comparativa associados a abordagens de genômica funcional e de genética molecular têm possibilitado o estudo detalhado de mecanismos potencialmente relevantes tanto para a colonização de plantas e insetos por este fitopatógeno como para o desenvolvimento de sintomas associados a doenças específicas em seus respectivos hospedeiros vegetais. Exceto o genoma de 9a5c, todos os demais genomas conhecidos são de cepas isoladas na América do Norte. Neste trabalho descrevemos o pirossequenciamento dos genomas da cepa J1a12, que exibe fenótipo não-virulento em citros, e das cepas Pr8x e Hib4, isoladas, respectivamente, de ameixeira e hibisco. A cepa J1a12 possui além de seu cromossomo principal de 2.788.789 pb dois plasmídeos, pXF51 e pXF27, respectivamente de 51.180 pb e 27.268 pb. pXF51 já foi descrito também na cepa de citros 9a5c e pXF27 tem similaridade com outros plasmídeos de cepas de *X. fastidiosa* norte-americanas isoladas de amoreira e videira. A cepa Pr8x possui além de seu cromossomo principal de 2.666.242 pb um plasmídeo, pXF39, de 39.580 pb, o qual contém a maioria das CDS presentes no pXF51. A cepa Hib4, isolada de hibisco, tem o maior cromossomo (2.813.297 pb) e também o maior plasmídeo (pXF64 com 64.251 pb) já descritos para *X. fastidiosa*. pXF64 apresenta extensa similaridade com o plasmídeo pBVIE04 de *Burkholderia vietnamsis* cepa G4, sendo descrito pela primeira vez em cepas de *X. fastidiosa*. Análises comparativas destes genomas possibilitaram a identificação de alterações que podem ser correlacionadas com os fenótipos exibidos por estas cepas, além da variedade e diversidade de regiões relacionadas a bacteriófagos e de plasmídeos que co-existem nas diferentes cepas deste fitopatógeno.

Palavras chave: Fitopatógeno, *Xylella fastidiosa*, Clorose Variegada dos Citros, Pirossequenciamento de Genomas; Genômica comparativa.

ABSTRACT

Pierry, Paulo Marques. **Pyrosequencing and comparative analysis of *Xylella fastidiosa* genomes**. 2012. 144p. Masters Thesis - Graduate Program in Biochemistry. Instituto de Química, Universidade de São Paulo, São Paulo.

Xylella fastidiosa is a Gram-negative bacteria, of the Gamma-proteobacterium subgroup, non-flagellated that colonizes the xylem of several cultivated and wild plants, where may cause disease. The bacterium is spread by insects known as sharpshooters. Genomes of *X. fastidiosa* strains isolated from different hosts have been completely or partially sequenced: 9a5c from citrus; Temecula-1 and GB514 from grapevine; Dixon, M12 and M23 from almond tree; Ann-1 from oleander and EB92-1, isolated from elderberry and used as bio-control for Pierce's disease of grapevines. Comparative genomics studies associated with approaches from functional genomics and molecular genetics have allowed a detailed study of mechanisms potentially relevant to the colonization of plants and insects by this pathogen as well as to the development of symptoms associated with specific diseases in their respective host plants. Except for 9a5c, all other known genomes are from strains isolated in North America. Here we describe the pyrosequencing of the genomes of strain J1a12, which displays non-virulent phenotype in citrus and of Pr8x and Hib4 strains isolated, respectively, from plum and hibiscus. J1a12 has a main chromosome of 2,788,789 bp and two plasmids, pXF51 and pXF27, respectively of 51,180 bp and 27,268 bp. pXF51 has been described also in the citrus strain 9a5c and pXF27 has similarity with other plasmids found in North American strains isolated from mulberry tree and grapevine. The strain Pr8x has a main chromosome of 2,666,242 bp and one plasmid, pXF39, of 39,580 bp which present similarities with pXF51. Hib4, the strain isolated from hibiscus, has the largest chromosome (2,813,297 bp) and the largest plasmid (pXF64 with 64,251 bp) described for *X. fastidiosa*. pXF64 shows extensive similarity with the plasmid pBVIE04 of *Burkholderia vietnamensis* G4 strain and is described for the first time in *X. fastidiosa*. Comparative analyzes of these genomes have identified several differences that may be correlated with the phenotypes displayed by these strains, in addition to the variety and diversity of regions related to bacteriophages and plasmids that co-exist in different strains of this pathogen.

Keywords: phytopathogen; *Xylella fastidiosa*; Citrus Variegated Chlorosis; genome pyrosequencing; comparative genomics.

LISTA DE ABREVIATURAS E SIGLAS

ALS: escaldadura da folha de amendoeira
BLAST: *Basic Local Alignment Search Tool*
CATG: Centro Avançado de Tecnologias em Genômica, IQUSP
CDS: sequência codificadora
CLS: Escaldadura da folha de cafeeiro
CTAB: brometo de n-hexadecil trimetil amônio
CVC: Clorose Variegada dos Citros
dNTP: desoxirribonucleotídeo 5' trifosfato
DO: densidade óptica
dsDNA: DNA fita-dupla
DSF: *Diffusible Signaling Factor*
EDTA: ácido etilenodiaminotetracético
emPCR: PCR em emulsão
Fundecitrus: Fundo de Defesa da Citricultura
IMG/ER: *Integrated Microbial Genomes-Expert Review*
kb: quilobase
LNCC: Laboratório Nacional de Computação Científica
MIRA: *Mimicking Intelligent Read Assembly*
MLSA: Análise de Sequência Multilocus
NCBI: *National Center for Biotechnology Information*
ORF: *open read frame*
pb: pares de base
PCR: reação em cadeia da polimerase
PD: Doença de Pierce
PET: *Paired End Tags*
PLS:escaldadura da folha de ameixeira
PTP: *PicoTiterPlate*
qPCR: PCR quantitativa
RAST: *Rapid Annotation using Subsystem Technology*
Rnase: Ribonuclease A
rpm: rotações por minuto
ssDNA: DNA fita-simples
TE: Tampão Tris contendo EDTA
T_m: temperatura de fusão
Tris: tris-(hidroximetil)-aminometano
UV: ultravioleta

SUMÁRIO

1. Introdução	12
1.1. O fitopatógeno <i>Xylella fastidiosa</i>	12
1.1.1. Variabilidade das linhagens de <i>X. fastidiosa</i>	15
1.1.2. Doenças causadas por <i>X. fastidiosa</i>	17
1.1.2.1. Clorose Variegada dos Citros (CVC).....	18
1.1.2.2. Escaldadura da folha de Ameixeira	20
1.1.2.3. Escaldadura de folha de Hibisco	22
1.2. Genomas de <i>X. fastidiosa</i>	22
1.3. Fatores de virulência e patogenicidade em <i>X. fastidiosa</i>	29
1.4. Estratégias de controle das doenças causadas por <i>Xylella</i>	34
2. Objetivos	38
3. Procedimentos experimentais	39
3.1. Manutenção e cultivo de <i>Xylella fastidiosa</i>	39
3.2. Avaliação morfológica de colônias de <i>X. fastidiosa</i>	39
3.3. Isolamento de DNA total	40
3.4. Isolamento de DNA plasmidial	41
3.5. Reação em cadeia da polimerase (PCR)	41
3.5.1. Genotipagem das cepas	41
3.5.2. Análise de pXF27 e pXF51	42
3.6. Eletroforese de DNA em gel de agarose	43
3.6.1. Eletroforese convencional	43
3.6.2. Eletroforese longa em baixa voltagem.....	44
3.7. Digestão do pXF51 e pXF27 com enzimas de restrição	44
3.7.1. Digestão do produto de PCR pXF27_circ	44
3.7.2. Digestão do DNA plasmidial total	45
3.8. Pirosequenciamento de DNA	45
3.8.1. Pirosequenciamento com a metodologia <i>Paired End Tags</i>	47
3.8.2. Pirosequenciamento com a metodologia <i>shotgun</i>	47
3.8.2.1. Reagentes utilizados	47
3.8.2.2. Quantificação do DNA genômico	48
3.8.2.3. Análise qualitativa do DNA genômico	48
3.8.2.4. Construção da biblioteca de DNA fita-simples	49
3.8.2.5. Titulação da biblioteca e PCR em emulsão em pequena escala.....	51
3.8.2.6. emPCR em grande escala ou definitiva	52
3.8.2.7. Preparação da placa e pirosequenciamento.....	53
3.8.2.8. Obtenção e processamento dos dados brutos	55
3.9. Análises bioinformáticas.....	55
3.9.1. Montagem e anotação de genomas	55
3.9.2. Análises filogenéticas.....	56
3.9.3. Outras análises bioinformáticas.....	57
3.10. Ensaios de infecção em cafeeiros com <i>X. fastidiosa</i>	57
3.10.1. Inoculação da suspensão bacteriana.....	57
3.10.2 Extração de DNA do pecíolo de plantas inoculadas com <i>X. fastidiosa</i>	59
3.10.3. Recuperação e isolamento de <i>X. fastidiosa</i> de plantas de café.....	60
4. Resultados e Discussão	62
4.1. Avaliação fenotípica de <i>Xylella fastidiosa</i> cepas J1a12, Pr8x e Hib4	62
4.2. Genotipagem de diferentes cepas de <i>X. fastidiosa</i>	64
4.3. Sequenciamento do genoma da cepa J1a12	66
4.3.1. Características do genoma da cepa J1a12.....	68
4.3.2. Análises comparativas do genoma de J1a12.....	73
4.4. Caracterização dos plasmídeos de pXF51 e pXF27 da cepa J1a12.....	88
4.4.1. Análises comparativas de pXF51.....	88

4.4.2. Análises comparativas de pXF27.....	88
4.4.3. Ensaios para identificação de pXF51 e pXF27 em <i>X. fastidiosa</i>	95
4.4.4. Análise de pXF27 e pXF51 de J1a12 com enzimas de restrição	98
4.4.5. Análise da presença pXF27 em outros isolados de <i>X. fastidiosa</i>	100
4.5. Sequenciamento do genoma das cepas Pr8x e Hib4	103
4.5.1. Montagem dos genomas de Pr8x e Hib4	106
4.5.2. Análises comparativas dos genomas de Pr8x e Hib4	110
4.5.2.1. Análises comparativas do genoma de Pr8x	110
4.5.2.2. Análises comparativas de pXF39.....	113
4.5.2.3. Análises comparativas do genoma de Hib4.....	115
4.5.2.4. Análises comparativas de pXF64.....	118
4.6. Análises filogenéticas de cepas de <i>X. fastidiosa</i>	120
4.7. Ensaio de infecção da cepa J1a12 em plantas de cafeeiro	123
5. Conclusões	125
6. Bibliografia	127
7. Apêndice	140
8. Súmula Curricular.....	145

1. Introdução

1.1. O fitopatógeno *Xylella fastidiosa*

O gênero *Xylella* é composto de uma única espécie, a *Xylella fastidiosa*. Esta é uma bactéria Gram-negativa, não flagelada, pertencente ao grupo das gamaproteobactérias e responsável por causar doenças nos mais variados hospedeiros vegetais, tais como a Doença de Pierce em videiras e a Clorose Variegada do Citros (Hopkins, 1989; Hopkins e Purcell, 2002; Chatterjee *et al.*, 2008; Janse e Obradovic, 2010).

Em 1884, uma doença em videira causou uma epidemia que destruiu 14.000 hectares (ha) de videiras (*Vitis* spp.) afetando dezenas de produtores de vinhos na área da bacia de Los Angeles nos Estados Unidos. Esta mesma doença foi também associada ao declínio da indústria de uva na Flórida nos anos 1890 (Nunney *et al.*, 2011). No ano 1892, o fitopatologista Newton B. Pierce descreveu esta doença que foi posteriormente chamada de Doença de Pierce (PD) (Hopkins e Purcell, 2002; Lauzière e Sétamou, 2009; Janse e Obradovic, 2010; Nunney *et al.*, 2011).

Até 1971, assumia-se que tanto PD como outra doença em pessegueiros (*phony peach disease*) eram causadas por vírus. Em 1978, Davis e colaboradores identificaram bactérias (*Xylella fastidiosa*) como agente etiológico dessas duas doenças (Davis *et al.*, 1978). Análises ultra-estruturais sugeriram que tal bactéria seria similar a bactérias das famílias *Rickettsiaceae* e *Legionellaceae*, sendo ocasionalmente referenciada como bactéria *rickettsia-like* (Wells *et al.*, 1987) e posteriormente como “bactéria limitada ao xilema” (XLB) (Hopkins, 1989). Comparações de sequências de rRNA 16S relacionaram estas bactérias fastidiosas ao grupo das Xanthomonadaceas, excluindo assim qualquer relação filogenética com as *rickettsiae* (Wells *et al.*, 1987; Hopkins, 1989).

As principais características que distinguem *X. fastidiosa* são seu confinamento no lúmen dos vasos xilemáticos de seus hospedeiros vegetais doentes, sua forma de bastonete com diâmetro entre 0,25 e 0,50 µm e comprimento entre 1,0 e 4,0 µm (pode formar filamentos de células em cultivo em determinados meios artificiais), colônias circulares opalescentes lisas ou rugosas, parede celular de bactérias Gram-negativas, positivas para catalase, aeróbicas obrigatórias, aflageladas, não-fermentativas, não-halofílica, não pigmentada e fastidiosas, com tempos de duplicação de 9 a 55 h dependendo do meio de cultivo. Temperaturas ótimas de crescimento são de 26 a 28°C, sendo que temperaturas abaixo de 10°C gradualmente reduzem, mas não eliminam populações da bactéria em videira. A

temperatura máxima de crescimento é de 33°C, a qual é surpreendentemente baixa para um patógeno que ocorre em climas tropicais e subtropicais. O valor de pH ótimo para o cultivo varia de 6,5 a 6,9, com valores de pH finais da cultura entre 7,2 e 7,8. A composição de ácido graxos, uma ferramenta taxonômica para distinguir entre gênero e possivelmente entre espécies de bactérias, foi similar para todas as linhagens testadas e apresenta uma alta porcentagem de ácidos graxos saturados (6 a 18%) com nenhum ácido ciclopropano (Davis *et al.*, 1978; Davis, 1981; Wells *et al.*, 1987; Hopkins, 1989; Hopkins e Purcell, 2002).

X. fastidiosa tem uma ampla abrangência de hospedeiros, causando doenças em mais de 100 espécies de plantas de pelo menos 46 famílias, contudo é considerada como patógeno fraco ou oportunista. Embora algumas linhagens pareçam ser restritas a um ou poucos hospedeiros, outras podem se multiplicar em vários. Muitas plantas selvagens podem carregar o patógeno em estado latente, sem mostrar sintomas, tais como gramíneas e plantas ornamentais, as quais servem como fontes de inóculos para insetos vetores. Além disso, ainda que todas as cepas já caracterizadas sejam fastidiosas e requeiram meios especializados para isolamento e crescimento, diferenças significantes nas taxas de crescimento são observadas entre linhagens isoladas de distintos hospedeiros (Wells *et al.*, 1987; Hopkins e Purcell, 2002; Schaad *et al.*, 2004; Hernandez-Martinez *et al.*, 2006; Hernandez-Martinez *et al.*, 2007; Lauzière e Sétamou, 2009; Janse e Obradovic, 2010).

A bactéria é transmitida tanto por insetos vetores da família Cicadellidae (*sharpshooters*) e Cercopidae (*spittlebugs*), popularmente conhecidos como cigarrinhas, como também por enxertia. Estes insetos são sugadores de seiva xilemática e ocasionalmente adquirem as bactérias ao se alimentarem de plantas contaminadas, sendo que raramente danificam as plantas diretamente (Davis *et al.*, 1978; Purcell e Finlay, 1979; Hill e Purcell, 1997; Almeida *et al.*, 2005; Almeida e Purcell, 2006; Damsteegt *et al.*, 2006; Alves *et al.*, 2008).

Algumas espécies de cigarrinhas contaminam apenas duas de 100 árvores que se alimentam (Molina *et al.*, 2010) e, embora a transmissão de *X. fastidiosa* pelo vetor pareça ser uma questão relativamente simples de adesão e desprendimento da bactéria das partes bucais, observações experimentais mostram que a transmissão é um processo complexo (Almeida *et al.*, 2005). A transmissão de *X. fastidiosa* é o resultado de três eventos: aquisição da bactéria de uma planta fonte, adesão e retenção na porção anterior do tubo intestinal do vetor, destacamento e inoculação da bactéria em um novo hospedeiro. Além disso, uma transmissão bem sucedida

também inclui o desenvolvimento de uma infecção, ou seja, múltiplos ciclos de multiplicação da bactéria após a inoculação (Almeida *et al.*, 2005).

X. fastidiosa é o único fitopatógeno transmitido por artrópodes que é propagativo (multiplica-se em seu vetor) sem ser circulativo na hemolinfa do inseto, sendo que o grande número de bactérias pode acumular no intestino do vetor com o passar do tempo. A bactéria se multiplica e se adere ao canal alimentar (precibário), à câmara de bombeamento (cibário) e à entrada do esôfago, regiões da parte anterior do intestino do inseto vetor nas quais podem ser formados densos agregados de células, inseridas em uma matriz (provavelmente exopolissacarídeo) que lhes oferece proteção física impedindo de serem levadas pela alta velocidade do fluxo durante a alimentação do inseto, além de poder contribuir na captação de nutrientes do fluido xilemático. As bactérias são encontradas nos vetores aderidas através de um de seus pólos, o que maximiza a sua exposição às correntes de fluido ingeridos, mas também podem ser encontradas lateralmente aderidas a várias regiões no cibário (Hill e Purcell, 1995; Redak *et al.*, 2004; Alves *et al.*, 2008; Backus e Morgan, 2011).

Os insetos vetores adultos são capazes de transmitir *Xylella fastidiosa* imediatamente após sua aquisição, sem a necessidade de um período latente e sem que a bactéria tenha colonizado as partes bucais e intestinais anteriores do inseto. Os insetos se mantêm infecciosos até a sua morte, o que pode acontecer após vários meses, sendo que este período de infectividade prolongado do inseto adulto pode ser devido à multiplicação da bactéria em seu tubo digestivo (Hill e Purcell, 1995; Brlansky *et al.*, 2002; Redak *et al.*, 2004; Almeida *et al.*, 2005; Alves *et al.*, 2008).

O fluido xilemático tem as concentrações mais diluídas de carbono e nitrogênio do que qualquer outro tecido da planta e pode variar significativamente dependendo de vários fatores tais como a estação do ano, hora do dia, fertilização e irrigação. Os constituintes orgânicos primários no fluido xilemático são ácidos orgânicos, açúcares e aminoácidos, em particular glutamina e asparagina, os quais são essenciais para a nutrição dos insetos, uma vez que podem ser convertidos em uma ampla gama de outros aminoácidos (Redak *et al.*, 2004; Alves *et al.*, 2008).

Assim como no hospedeiro vegetal (Araújo *et al.*, 2002), também no inseto-vetor *X. fastidiosa* pode conviver em uma comunidade microbiana diversa que inclui diversos gêneros tais como *Chryseomonas*, *Ralstonia*, *Alcaligenes*, *Wolbachia*, *Bacillus*, *Pseudomonas*, *Pedobacter*, *Methylobacterium*, *Curtobacterium* e *Burkholderia* (Bextine *et al.*, 2004; Lacava *et al.*, 2007; Backus e Morgan, 2011; Hail *et al.*, 2011). A competição por sítios de ligação no aparelho bucal do inseto pode

explicar diferenças nas eficiências de aquisição de *X. fastidiosa* pelos insetos vetores, e consequentemente pela competência do vetor e eficiência de inoculação. Talvez a riqueza de espécies da comunidade microbiana habitante do xilema em hospedeiros vegetais varie entre as diferentes áreas geográficas, o que influencia a estrutura da comunidade microbiana no tubo digestivo do vetor (Backus e Morgan, 2011). Existem evidências que bactérias fitopatogênicas podem ter evoluído associações específicas e não-específicas com insetos, os quais elas exploram como veículos de transferência ou como hospedeiros primários alternativos (Nadarasah e Stavrinides, 2011).

1.1.1. Variabilidade das linhagens de *X. fastidiosa*

Espécies bacterianas são normalmente subdivididas pelo reconhecimento de sorovares, patovares, isolados, cepas ou linhagens que tipicamente infectam diferentes hospedeiros. Além disso, muitas espécies, assim como *X. fastidiosa*, exibem um alto grau de diversidade genética entre linhagens que infectam o mesmo hospedeiro, e este padrão é típico da subdivisão gama das proteobacterias (Schuenzel *et al.*, 2005; Studholme *et al.*, 2005; Williams *et al.*, 2010).

As linhagens de *Xylella fastidiosa* diferem o suficiente fenotipicamente e genotipicamente de outras Xanthomonadaceas para justificar a separação em gêneros diferentes. Porém, ainda que todas elas tenham sido classificadas como uma única espécie, há variabilidade suficiente para justificar separação taxonômica ao nível de subespécie ou patovar (Hopkins, 1989).

A variabilidade entre linhagens de *X. fastidiosa* é notada mesmo quando genes altamente conservados, tais como rDNA 16S são comparados. Uma árvore filogenética, construída utilizando sequências de rDNA 16S de linhagens de *X. fastidiosa*, separou dois grupos maiores. O grupo I abrangeu linhagens de videira e a linhagem de amoreira, enquanto que o grupo II incluiu linhagens isoladas de laranjeira, cafeeiro, ameixeira, vinca e pessegueiro. Dois outros grupos menores foram formados por linhagens únicas de carvalho e pêra (Mehta e Rosato, 2001).

Em outro estudo filogenético, foi construída uma árvore de máxima verossimilhança usando 25 isolados de *X. fastidiosa* da América do Norte, com a linhagem de CVC da América do Sul como grupo externo. Três clados maiores foram obtidos, sendo que dois correspondem a *X. fastidiosa* subsp. *piercei* (*fastidiosa*) e *X. fastidiosa* subsp. *multiplex*. O clado da subespécie *piercei* (*fastidiosa*) inclui todos os isolados de videira mais um subconjunto de isolados de amendoeira, e o clado da subespécie *multiplex* inclui isolado de pêsego, ameixeira,

e outro subconjunto de isolados de amendoeira, além de três isolados de carvalho. O terceiro clado agrupou todos os isolados de espiroleira. O ramo unindo este grupo com o da subespécie *piercei* (*fastidiosa*) está completamente suportado por *bootstrap*, colocando estes dois clados como clados irmãos, mais relacionados entre si do que com o clado da subespécie *multiplex*. Os autores então sugerem que os isolados de espiroleira devam ser reconhecidos como uma nova subespécie com o nome *X. fastidiosa* subsp. *sandyi*, sendo o isolado Ann-1 um dos representantes (Schuenzel *et al.*, 2005). Este mesmo estudo sugere que o ancestral de *X. fastidiosa* foi um organismo clonal que divergiu há mais de 15.000 anos nos três clados, *X. fastidiosa* subsp. *piercei* (*fastidiosa*), *multiplex* e *sandyi*. Esta inferência sugere a longa existência de linhagens de *X. fastidiosa* nas Américas antes da introdução de plantas não-nativas tais como videira e citros que exibem sintomas distintos a partir da infecção (Doddapaneni *et al.*, 2007).

Montero-Astúa e colaboradores (2007) analisaram através de diferentes métodos a diversidade de 42 linhagens *X. fastidiosa* dos Estados Unidos, Brasil e Costa Rica e concluíram que há uma similaridade genética maior entre as linhagens da Costa Rica e Estados Unidos do que com as linhagens isoladas no Brasil. Mostraram também que isolados de diferentes hospedeiros na Costa Rica são mais próximos entre si do que com isolados das outras regiões, refletindo o surgimento deste patógeno simultaneamente em regiões geográficas distintas nas Américas ou alternativamente a introdução recente de um ancestral da linhagem de Temecula1 na Costa Rica (Montero-Astúa *et al.*, 2007).

Em outro estudo baseado em filogenias do rRNA 16S, foi proposto que linhagens de *X. fastidiosa* isoladas da planta ornamental chitalpa (*Chitalpa tashkentensis*), comum em regiões áridas dos Estados Unidos, são distintas das subespécies anteriormente descritas, sendo então proposta a nova subespécie *X. fastidiosa* subsp. *tashke* (Randall *et al.*, 2009).

Em resumo, as linhagens de *X. fastidiosa* foram separadas, até o momento, nas seguintes subespécies (Schaad *et al.*, 2004; Schuenzel *et al.*, 2005; Hernandez-Martinez *et al.*, 2007; Randall *et al.*, 2009; Janse e Obradovic, 2010):

- *fastidiosa* ou *piercei*, para linhagens de videiras, alfalfa, amendoeira, e bordo;
- *multiplex*, para linhagens de pêssgo, olmo, ameixeira, sicômoro e amendoeira;
- *pauca*, para linhagens de citros, provavelmente de café e outras linhagens da América do Sul;

- *sandyi*, para linhagens de espiroleira, jacarandá, magnólia e *daylily*;
- *tashke*, para linhagens de chitalpa.

Além destas cinco subespécies, uma sexta subespécie foi proposta para linhagens de amoreira (*Morus* spp.) e bambú (*Nandina domestica*) (Hernandez-Martinez *et al.*, 2007). Estes autores também confirmaram que nenhuma das linhagens de diversos outros hospedeiros norte-americanos tem qualquer relação com subespécie *pauca*.

1.1.2. Doenças causadas por *X. fastidiosa*

A maioria das doenças causadas por *Xylella* ocorrem na América do Norte e são particularmente severas em regiões tropicais e subtropicais com invernos moderados, incluindo climas mediterrâneos. Porém doenças de escaldadura da folha também ocorram em climas mais frios, como por exemplo, a escaldadura da folha de carvalho encontrada até o Canadá. Algumas destas doenças norte-americanas têm uma distribuição mais ampla, tais como a Doença de Pierce, que também ocorre na América Central (Costa Rica), México e Venezuela, a escaldadura da folha de ameixeira (PLS) que ocorre também no Brasil, Paraguai e Argentina e escaldadura da folha de amendoeira (ALS) que é encontrada também na Argentina. Exceções são CVC (Clorose Variegada dos Citros) e CLS (escaldadura da folha de cafeeiro), as quais até o momento parecem ser restritas a América do Sul e provavelmente a América Central (Hopkins, 1989; Redak *et al.*, 2004; Janse e Obradovic, 2010).

Doenças causadas por *X. fastidiosa* são observadas, principalmente, em plantações que são exóticas às Américas (videira, amendoeira, alfalfa, pessegueiro, citros, café, e ameixeiras domésticas) enquanto que diversas espécies de plantas nativas selvagens são hospedeiros tolerantes (suportam a colonização bacteriana, mas não mostram sintomas), sugerindo que esta bactéria é endêmica neste continente há um longo tempo. Por outro lado, doenças causadas por linhagens de *X. fastidiosa* também ocorrem em árvores nativas em florestas norte-americanas, onde espécies de carvalho e sicômoro entre outras árvores são afetadas (Hopkins, 1989; Redak *et al.*, 2004; Janse e Obradovic, 2010).

X. fastidiosa foi, aparentemente, associada à escaldadura da folha de pereiras japonesa na Tailândia e na Austrália; relatos associaram bactérias pequenas, fastidiosas e limitadas ao xilema com a doença de Sumatra em cravos-da-Índia e doença do enrugamento da folha de *Hibiscus cannabinus* L. A

escaldadura da folha de amendoeira na Índia também foi associada à *Xylella* (Schaad *et al.*, 2004; Janse e Obradovic, 2010).

Interessantemente, doenças causadas por *Xylella* aparentemente estão ausentes na Europa, exceto por um relato da doença de Pierce em videiras na região da Sérvia, em 1980, sendo que a linhagem isolada mostrou o mesmo patotipo da linhagem causadora desta doença na América do Norte. Postula-se que a doença tenha sido introduzida por cultivares de videiras ou porta-enxertos norte-americanos contaminados, mas que não tenha sido disseminada devido às condições climáticas com temperaturas mais frias e a ausência de insetos vetores eficientes para a transmissão (Berisha *et al.*, 1998).

O rápido surgimento e a rápida disseminação da CVC no Brasil tiveram início no final da década de 1980 e alcançou todas as regiões citrícolas do Brasil em uma década (Rossetti *et al.*, 1990; Chang *et al.*, 1993; Coletta-Filho e Machado, 2003). A doença se espalhou rapidamente na América do Sul e foi identificada na Costa Rica em 2005 (Montero-Astúa *et al.*, 2007). Embora a linhagem associada à CVC não tenha sido encontrada nos Estados Unidos, não está descartada sua ameaça às produções de laranjas na América Central e do Norte se um vetor adequado estiver disponível (Redak *et al.*, 2004; Damsteegt *et al.*, 2006).

A origem da linhagem de *X. fastidiosa* causadora da doença de Pierce, ainda é alvo de discussão. A hipótese mais aceita é que um único genótipo de *X. fastidiosa* subsp. *fastidiosa*, causador da doença de Pierce, foi introduzido nos Estados Unidos algum tempo pouco antes do surto de Anaheim nos anos 1880s. *X. fastidiosa* subsp. *pauca* é altamente diferenciada das outras três subespécies (cerca de 3%), sugerindo isolamento a longo prazo, o que torna a América do Sul uma fonte improvável para *X. fastidiosa* subsp. *fastidiosa*. Assim, uma fonte mais provável seria a América Central. Como já mencionado, isolados da Costa Rica foram mais similares geneticamente a isolados dos Estados Unidos do que com os do Brasil. Muitas plantações diferentes foram importadas para a Califórnia no período de 1850-1870, inclusive plantas de cafeeiro no sul da Califórnia oriundas da América Central (Nicarágua) o que sugere sua introdução na América do Norte tenha ocorrido através de outras plantas hospedeiro (Nunney *et al.*, 2011).

1.1.2.1. Clorose Variegada dos Citros (CVC)

A doença causada pela bactéria *Xylella fastidiosa* em espécies de citros foi encontrada pela primeira vez no Brasil em 1987 em pomares do município de Colina, Estado de São Paulo, e na região do Estado de Minas Gerais (Rossetti e De

Negri, 1990; Coletta-Filho e Machado, 2003). No Estado do Paraná, a CVC foi observada, primeiramente, na cidade de Paranavaí, em 1992 (Molina *et al.*, 2010). *X. fastidiosa* apresenta uma grande disseminação e ampla distribuição em culturas espalhadas por todo país contrastando com a ocorrência de *Xanthomonas axonopodis* pv. *citri*, bactéria causadora do cancro cítrico, que é limitada às regiões Sul e Sudeste (Zambolim e Bassanezi, 2006). Na Argentina, a CVC foi detectada pela primeira vez em 1984, onde foi chamada de *Pecosita* (Hopkins e Purcell, 2002). Também foi relatada no Paraguai e Costa Rica (Coletta-Filho e Machado, 2003; Montero-Astúa *et al.*, 2007).

Entre as espécies de citros que podem ser infectadas, tanto cultivares e híbridos de *Citrus* spp, as laranjeiras doce (*Citrus sinensis*) são as mais susceptíveis. Em mandarins (*Citrus reticulata*) os sintomas revelados são menos severos. Limão-cravo (*Citrus limonia*), limão (*Citrus limon*), cidra (*Citrus medica*) e pumelo (*Citrus grandis*) são tolerantes para a doença (Rossetti e De Negri, 1990). A CVC não mata as plantas susceptíveis, mas causa danos à cultura diminuindo a qualidade das frutas, tornando-as inadequadas para a comercialização “*in natura*”, assim como para a produção de sucos concentrados (Li *et al.*, 2003; Bove e Ayres, 2007; Molina *et al.*, 2010).

Embora *X. fastidiosa* tenha sido encontrada em pedúnculos de laranja doce quando a bactéria foi pela primeira vez associada com a doença, o fruto e as sementes também são afetados e podem abrigar o patógeno. *X. fastidiosa* foi detectada por PCR em amostras do pedúnculo, casca, endocarpo, septo, membrana locular, e eixo central de frutas afetadas por CVC em três cultivares analisados (Pêra, Natal e Valencia). A bactéria foi também detectada nas sementes da laranja, as quais são abundantemente dotadas de vasos do xilema que são diretamente conectados ao sistema vascular da planta (Li *et al.*, 2003).

A deposição de goma nos vasos é um dos primeiros sintomas internos, além da formação de tiloses. Entre os primeiros sintomas externos da doença estão lesões cloróticas foliares, principalmente intervenais nas nervuras das folhas, muito similar aos sintomas de deficiência de zinco. Em estágios posteriores, pontos necróticos marrons se desenvolvem na parte de baixo da folha, correspondendo a áreas cloróticas na parte de cima, e, além disso, os frutos se tornam pequenos e endurecidos. Déficit prolongado de água pode exacerbar o desenvolvimento de sintomas da doença (Hopkins, 1989; Janse e Obradovic, 2010).

Por muito tempo, o surgimento dos sintomas da doença foi associado a oclusão e entupimento dos vasos do xilema, resultando em estresse hídrico nos

ramos e folhas superiores ao local da inoculação da bactéria (Hopkins, 1989; Purcell e Hopkins, 1996). Entretanto, estudos realizados com videiras desassociaram os sintomas da oclusão dos vasos, mostrando que os sintomas de estresse hídrico e da Doença de Pierce eram distintos (Thorne *et al.*, 2006). Além disso, parece existir pouca ou nenhuma correlação entre concentrações localizadas da bactéria e desenvolvimento de sintomas, sugerindo que altas concentrações localizadas de *X. fastidiosa* não são necessárias para o aparecimento de sintomas de escaldadura da folha e que respostas da planta podem estar associadas à patogênese em videiras (Gambetta *et al.*, 2007). Além disso, a estrutura bruta do xilema e sua organização não são, aparentemente, responsáveis por diferenças genotípicas na susceptibilidade da planta à infecção por *X. fastidiosa*, embora diferenças estruturais da membrana do poro das pontuações nos vasos do xilema poderiam responder por diferenças entre tolerância e susceptibilidade à doença (Thorne *et al.*, 2006; Pérez-Donoso *et al.*, 2007; Chatelet *et al.*, 2011).

Também em citros, outras hipóteses têm sido propostas para explicar o desenvolvimento de sintomas da CVC, incluindo a produção de fitotoxinas pelos patógenos e sequestro de nutrientes, particularmente cálcio, magnésio e zinco, atraídos pela alta densidade de cargas negativas nos agregados bacterianos (Hopkins, 1989; Simpson *et al.*, 2000; Leite *et al.*, 2002). Em um estudo com plantas de café, laranja e ameixa, folhas com sintomas moderados de todos os hospedeiros foram frequentemente encontradas em uma porção mais distal do ramo, enquanto que as folhas com sintomas severos foram encontradas mais próximas do caule principal. A porcentagem de vasos do xilema colonizados e ocluídos foi maior em folhas em cafeeiro, seguido de ameixeira e citros, sugerindo que a porcentagem de vasos do xilema colonizados é influenciada pelo hospedeiro vegetal (Alves *et al.*, 2004).

1.1.2.2. Escaldadura da folha de Ameixeira

A ameixeira é uma das árvores frutíferas de cultivo mais antigo no Brasil, não havendo informações precisas sobre quando foi introduzida no país. Pertence à família das Rosáceas (subfamília Prunoideae) e ao gênero *Prunus* (L.), sendo as espécies *Prunus salicina* Lindl. e *Prunus domestica* L. com maior importância econômica. A produção mundial de ameixa é de aproximadamente seis milhões de toneladas, destinada ao consumo *in natura*, passa e à elaboração de destilados. O mercado brasileiro é de cerca de 50 mil toneladas anuais, dos quais aproximadamente 50% são importadas principalmente do Chile e da Argentina. Os

principais pólos produtores de ameixa no Brasil são os estados do Rio Grande do Sul, Santa Catarina, São Paulo, Paraná e Minas Gerais (revisado por (Muller, 2009).

A escaldadura da folha de ameixeira (*Plum Leaf Scald* – PLS), doença causada pela bactéria *X. fastidiosa*, foi relatada pela primeira vez pela na região do delta do rio Paraná na Argentina por volta de 1935 (Kitajima *et al.*, 1975). No Brasil, a PLS foi inicialmente observada em 1978 em Cascata, no estado do Rio Grande do Sul, em árvores de ameixeira japonesa (*Prunus salicina* Lindl.) e européia (*Prunus domestica* L.) (French e Kitajima, 1978), embora produtores indiquem que a doença estava presente no Estado nos anos 50. PLS foi amplamente disseminada no sul do Brasil e também no Paraguai. Nos Estados Unidos, a PLS foi relatada em 1977 (French e Feliciano, 1982; Raju *et al.*, 1982). Hoje, a PLS está presente em todos os estados do sul do Brasil onde ameixeira é crescida comercialmente (Lopes *et al.*, 2003). O emprego de cultivares tolerantes tem sido o melhor método de controle da doença (Kitajima *et al.*, 1975).

Sintomas normalmente surgem no final de dezembro e começo de janeiro e ocorrem inicialmente em folhas antigas. O primeiro sintoma é uma leve e irregular clorose na borda da folha, tanto na margem como na ponta, intensificando-se mais tarde. Os sintomas aparecem primeiro nas folhas mais baixas do ramo e avançam progressivamente em direção a ponta. Entre os sintomas tardios se observam as margens da folha marrons e secas, sendo separadas de áreas não afetadas por uma banda difusa de clorose. A escaldadura pode aparecer em uma ou mais áreas de uma folha afetada e pode envolver três quartos da folha antes de abscisão ocorrer. Os ramos superiores tornam-se fracos e quebradiços, e em fases mais avançadas da doença, as plantas param de crescer e morrem em alguns anos. Os frutos, em cultivares altamente susceptíveis, não atingem tamanhos normais, e a produção é gradualmente reduzida com o progresso da doença. A PLS pode ser transmitida por enxertia e os primeiros sintomas aparecem em plantas inoculadas após meses ou anos de infecção (Kitajima *et al.*, 1975; Latham e Norton, 1980; French e Feliciano, 1982; Raju *et al.*, 1982).

Raju e colaboradores (1982) realizaram pela primeira vez o cultivo da bactéria associada com PLS, analisaram suas características morfológicas e demonstraram a proximidade filogenética com a bactéria responsável pela doença “*phony peach*” em pessegueiros (Raju *et al.*, 1982).

Hernandez-Martinez e colaboradores (2009) utilizaram isolados de *X. fastidiosa* obtidos da espécie *Prunus cerasifera* na Califórnia para análises de perfis de RAPD-PCR e sequências de 16S-23S rDNA. As linhagens extraídas de *P.*

cerasifera mostraram pertencer ao grupo da subespécie *X. fastidiosa* subsp. *multiplex*. Uma linhagem de *P. cerasifera* mostrou causar sintomas da doença quando inoculadas em ameixeira e amendoeira (subespécie *multiplex*), mas não gerou sintomas em espiroleira (subespécie *sandyi*) e videira (subespécie *fastidiosa*) (Hernandez-Martinez *et al.*, 2009).

1.1.2.3. Escaldadura de folha de Hibisco

Poucas informações existem a respeito da escaldadura da folha de hibisco e da caracterização de seus isolados. Uma das únicas linhagens de hibisco até então relatada foi isolada no Centro de Citricultura Sylvio Moreira (Cordeirópolis, SP) partir de plantas coletadas em Brasília (Kitajima *et al.*, 2000; Coletta-Filho *et al.*, 2001). Plantas deste gênero, identificadas como um reservatório potencial de *X.fastidiosa* (McGaha *et al.*, 2007), já foram utilizadas em estudos de transmissão do fitopatógeno pelos seus vetores (Lauzière e Sétamou, 2009) e de caracterização do conteúdo de fagos bacterianos em linhagens de *X. fastidiosa* (Chen e Civerolo, 2008).

Lauzière e Sétamou (2009) mostraram que hibisco (*Hibiscus rosa-sinensis*) foi a planta preferencial, entre os hospedeiros utilizados, para ovoposição do inseto vetor. Estes autores relataram que espécies de *Hibiscus* spp são visitadas pelas fêmeas para aquisição de nutrientes, especificamente nos meses de verão (Lauzière e Sétamou, 2009).

Não há um estudo detalhado sobre a manifestação de sintomas em hibisco.

1.2. Genomas de *X. fastidiosa*

X. fastidiosa foi o primeiro fitopatógeno bacteriano a ter o seu genoma completamente sequenciado e publicado (Simpson *et al.*, 2000). A linhagem escolhida foi a 9a5c, um isolado obtido de laranja doce (*Citrus sinensis*) do município de Macaúbal no estado de São Paulo (Li *et al.*, 1999). Esta linhagem de *X. fastidiosa* possui um cromossomo principal com tamanho de 2.679.306 pb, além de dois plasmídeos, pXF51 (51.158 pb) e pXF1.3 (1.286 pb). Foram preditas 2.782 ORFs no cromossomo principal, sendo que para 58.4% foram designadas funções putativas baseado em homologia com genes conhecidos; para o pXF51 foram preditas 61 ORFs e duas para o pXF1.3 (Simpson *et al.*, 2000; Marques *et al.*, 2001).

Em 2002, as sequências incompletas de dois genomas de *X.fastidiosa* foram divulgadas com “*draft genomes*”. Os isolados sequenciados foram a linhagem Dixon,

obtida de amendoeira e, a linhagem Ann-1, obtida de espiroleira, ambas de plantas com sintomas de escaldadura da folha e obtidas da coleção de isolados de Alexander H. Purcell, da Universidade da Califórnia, Berkeley, EUA. Foi proposto um cromossomo de 2.629.797 pb, além de um plasmídeo de 31.572 pb para a linhagem Dixon e um cromossomo de 2.617.604 pb, além de um plasmídeo de 30.270 pb para a linhagem Ann-1 (Bhattacharyya *et al.*, 2002). Mais tarde se verificou que a preparação de DNA utilizada como sendo de uma única cepa (Ann-1) estava contaminada com DNA de outra cepa não identificada, e que, portanto as sequências descritas para Ann-1 correspondem de fato a duas cepas (Nunney *et al.*, 2012).

Em 2003, foi publicada a sequência completa do genoma de uma linhagem de *X. fastidiosa*, isolada de videira com sintomas de Doença de Pierce, denominada Temecula1, obtida de uma região homônima produtora de uva na Califórnia, EUA, em 1998. A linhagem Temecula1 de *X. fastidiosa* tem um cromossomo principal com tamanho de 2.519.802 pb, além de um plasmídeo de 1.346 pb. Um total de 2.066 ORFs foi predito no cromossomo principal, das quais 65,9% foram designadas funções putativas baseado em homologia com genes conhecidos (Van Sluys *et al.*, 2003). Neste trabalho, os autores revisaram a anotação originalmente publicada para a cepa 9a5c, reduzindo o número de ORFs preditas. Recentemente, foi divulgado que a cepa Temecula1 sequenciada não tem correspondência com a cepa distribuída como “Temecula1” pela ATCC (*American Type Tissue Collection*), indicando assim a importância da genotipagem para verificação da autenticidade de cepas rotineiramente utilizadas (Nunney *et al.*, 2012).

Em 2010, foram publicadas as sequências completas dos genomas de outras duas linhagens de *X. fastidiosa*, a linhagem M12 e M23, ambas isoladas de amendoeiras com sintomas de escaldadura da folha de amendoeira, obtidas em Kern County em San Joaquin Valley na Califórnia, Estados Unidos. A linhagem M23 também pode causar a Doença de Pierce em videiras. A linhagem M12 de *X. fastidiosa* tem um cromossomo principal com tamanho de 2.475.130 pb, e nenhum plasmídeo foi encontrado. Um total de 2.368 ORFs foi predito no cromossomo principal. A linhagem M23 de *X. fastidiosa* tem um cromossomo principal com tamanho de 2.535.690 pb, e um plasmídeo de 38.297 pb. Um total de 2.280 ORFs foi predito no cromossomo principal (Chen *et al.*, 2010).

Também em 2010, a primeira sequência de um genoma de uma linhagem de *X. fastidiosa* do estado do Texas (Estados Unidos) foi publicada. A linhagem, denominada GilBec 514 ou GB514, foi obtida de uma videira com sintomas de

Doença de Pierce em uma região produtora em Gillespie County, no Texas. Esta linhagem de *X. fastidiosa* tem um cromossomo principal com tamanho de 2.491.203 pb, e um plasmídeo de 26.180 pb. Assim como os genomas das linhagens Dixon e Ann-1, este genoma também é um “draft”, ou seja, uma sequência incompleta do genoma (Schreiber *et al.*, 2010).

A última sequência de um genoma de *X. fastidiosa* foi publicada em 2011, e corresponde a uma linhagem isolada de *elderberry* (sabugueiro do Canadá) denominada EB92-1, obtida por D. L. Hopkins em 1992. Esta linhagem infecta videiras, consegue colonizar e sobreviver por muitos anos, mas não causa sintomas e promove um efetivo biocontrole contra a Doença de Pierce. Também corresponde a um “draft” com a sequência incompleta do genoma somando 2.478.730 pb e um plasmídeo de 1.346 pb. Apesar da alta similaridade e sintênia dos genomas de EB92-1 e Temecula1, foi detectada a ausência de 10 genes na cepa EB92-1 que podem, eventualmente, explicar seu fenótipo não-virulento. Entre os genes ausentes, estão CDSs que codificam uma serina-protease e uma lipase secretadas, dois genes que codificam a toxina de zonula occludens (Zot) além de seis genes de hemaglutininas (Zhang *et al.*, 2011).

O genoma de *X. fastidiosa* tem praticamente a metade do tamanho do genoma de bactérias do gênero *Xanthomonas* (~5,5 Mpb) o que provavelmente reflete suas estratégias diferentes de colonização, virulência e patogenicidade. Enquanto *X. fastidiosa* depende das cigarrinhas para a sua transmissão sendo limitada ao xilema em planta, *Xanthomonas* pode infectar plantas intactas e é capaz de crescer e sobreviver nas superfícies das folhas assim como no solo. O tamanho relativo dos genomas também indica que existem diferenças no complemento de genes carregados pelas duas bactérias, embora partilhem muitos genes ortólogos. A principal diferença entre *Xylella* e *Xanthomonas* é a ausência, no genoma de *X. fastidiosa*, de genes que codificam o sistema de secreção do tipo III. Sistemas de secreção do tipo III são amplamente distribuídos em bactérias fitopatogênicas, onde eles funcionam para secretar efetores do patógeno diretamente no interior das células da planta. Alguns destes determinantes de virulência podem atuar para suprimir as respostas de defesa da planta e dessa forma promover a doença, enquanto que outros podem atuar para promover a liberação de nutrientes da célula vegetal. Em alguns casos, entretanto, efetores secretados podem ser reconhecidos pela planta e ativar a defesa da planta, sendo considerados como determinantes de avirulência. Entretanto, o sistema de secreção do tipo II está completo em *X. fastidiosa* e parece ser responsável por exportar

enzimas degradadoras de parede celular presentes no genoma deste fitopatógeno (Dow e Daniels, 2000; da Silva *et al.*, 2002; Moreira *et al.*, 2005).

O genoma da cepa 9a5c de *X. fastidiosa* compartilha mais de 1400 CDSs com o genoma de *Xanthomonas axonopodis citri* (XAC), agente causal do cancro cítrico. Apesar de ter um genoma com tamanho muito menor, *Xylella* tem genes que não são encontrados em XAC. Por outro lado, XAC tem 6 cópias de genes para enzimas pectinolíticas e 12 cópias de genes para enzimas celulolíticas e hemicelulolíticas, enquanto que o genoma de 9a5c codifica apenas uma enzima pectinolítica e 4 enzimas celulolíticas. Os gêneros *Xanthomonas* e *Xylella* são os únicos organismos que contêm o *cluster rpf*. Porém, um rearranjo no genoma de *Xylella* separa os genes *rpfB* e *rpfF*, relacionados com a síntese e processamento de DSF (*Diffusible Signaling Factor*), dividindo o *cluster* em duas partes com a deleção de pelo menos 3 genes desse *cluster* (da Silva *et al.*, 2002; Moreira *et al.*, 2004; Moreira *et al.*, 2005; Chatterjee *et al.*, 2008).

A comparação das sequências dos genomas das linhagens 9a5c, Dixon e Ann-1 mostrou que estas linhagens conservam entre si mais de 85% de suas ORFs e que as diferenças concentram-se em regiões associadas a fagos (Bhattacharyya *et al.*, 2002). Análises comparativas da sequência e anotações completas dos genomas das cepas 9a5c e Temecula1 também revelaram uma variabilidade genômica limitada, sendo que estas cepas compartilham 95,7% de identidade de aminoácido em regiões equivalentes. Foram encontrados 41 genes específicos de Temecula1 e 152 de 9a5c e, uma diferença de 159.503 pb no tamanho do cromossomo principal. Ambas contêm um plasmídeo em comum de ~1,3 kpb, embora o plasmídeo pXF51 esteja ausente em Temecula1. Dos 2.066 genes codificantes de proteínas anotados em Temecula1, 2.025 (98%) também estão presentes em 9a5c. O total de 106 CDS no genoma de Temecula1 (5.2%), embora presentes no genoma de 9a5c, tem identidade de sequência de aminoácidos menor que 80%. Entre estes estão 58 genes que são encontrados dentro de regiões relacionadas a fagos e ilhas genômicas. Interessantemente, entre os genes com funções definidas com um nível maior de divergência, foram encontrados alguns que podem estar envolvidos em interações planta-hospedeiro, incluindo adesinas, fimbrilinas, hemaglutininas, colicinas, hemolisinas, bacteriocinas, proteínas de resistência a drogas e enzimas de restrição e modificação de DNA. Com o alinhamento dos cromossomos de Temecula1 e 9a5c, se destacaram três regiões cromossômicas que foram translocadas e invertidas apesar de sua identidade geral. Se as regiões de prófagos forem excluídas e rearranjos forem reorientados, os genomas de ambas as

linhagens de *X. fastidiosa* seriam muito similares e colineares (Van Sluys *et al.*, 2003).

Análises genômicas comparativas das linhagens 9a5c, Temecula, Ann-1 e Dixon apontaram 1.579 genes homólogos representando 87 famílias e 194 sequências não-codificantes, as quais correspondem a, respectivamente, 72,6% e 3,6% do tamanho total destes genomas. Baseado em análises por buscas por BLAST, 108 CDS (6.8%) foram únicas de *X. fastidiosa* com nenhuma similaridade com genomas bacterianos sequenciados até então. O número mais alto de genes específicos de linhagem foi para 9a5c (241 genes) seguida pela Ann1 (145 genes), Dixon (96 genes) e Temecula-1 (10 genes) (Doddapaneni *et al.*, 2007). Embora, estes autores não tenham considerado eventuais diferenças nos critérios utilizados na anotação destes genomas, foi sugerido que o menor número de genes específicos em Temecula1 sugere que ela possa estar mais próxima do genoma ancestral de *X. fastidiosa*. Além disso, alta similaridade entre os genomas de 9a5c e Temecula1 poderia indicar a origem recente de 9a5c a partir de Temecula1. Devido a um maior número de genes específicos em 9a5c, pode-se sugerir que ela esteja evoluindo em uma taxa mais rápida comparada com as outras três linhagens. Todos estes estudos comparativos citados sugerem que as diferenças em patogenicidade entre as linhagens de *X. fastidiosa* e na especificidade de hospedeiro são geneticamente controladas. *X. fastidiosa* não contém regiões de mono- e di-repetições em seu genoma mas é rico em repetições de 7- e 8 nucleotídeos. Estes resultados estão em contraste com outras bactérias Gram-negativas tais como *E. coli*, como a linhagem K12 que apresenta 19.200 mono-repetições e 7.575 di-repetições (Doddapaneni *et al.*, 2007).

Além de análises comparativas das sequências dos genomas, estudos utilizando hibridização de microarranjos de DNA compararam várias linhagens de *X. fastidiosa*. Em um primeiro estudo, sequências representativas de 2200 CDS do genoma de 9a5c foram hibridizadas com DNA de 11 linhagens de *Xylella*, obtidas de diferentes hospedeiros vegetais e origens geográficas. Os resultados revelaram a existência de um conjunto principal de genes altamente conservados, contendo todos os genes envolvidos nas principais vias bioquímicas e funções celulares básicas, as quais possivelmente determinam a adaptação desta bactéria ao crescimento no xilema da planta. Além disso, foi encontrado um conjunto de genes que codificam funções adicionais não essenciais para o crescimento bacteriano, mas que possivelmente confere vantagens sob condições particulares, tais como mudanças no ambiente e adaptação a um novo hospedeiro. Estes genes são

normalmente associados com fagos, plasmídeos, integrons e outros elementos transponíveis (Nunes *et al.*, 2003).

Outro estudo envolvendo hibridização de microarranjos de DNA foi realizado com seis linhagens sul-americanas de *Xylella*, todas obtidas de árvores de laranjeira infectadas: quatro (56a, 9.12c, 187b, e 36f) foram obtidas de árvores afetadas por CVC e são representativas de haplótipos prevalentes encontrados em pomares de laranjeira doce ao longo do estado de São Paulo; linhagem Cv21, obtida de uma árvore não-sintomática da mesma região; e linhagem Fb7, obtida de uma laranjeira doce que mostrou sintomas de "*Pecosita*", uma doença similar a CVC que ocorre nas regiões citrícolas da Argentina. A maioria das diferenças genômicas dentro de linhagens de citros não estão associadas com ORFs deletadas, mas com elementos que estão presentes em maior número de cópias nas linhagens testadas, quando comparadas com a linhagem 9a5c. Como esperado, cada linhagem apresentou um perfil genômico único, o qual pode ser usado para caracterizar todas as linhagens individuais com alta fidelidade. A linhagem da Argentina, Fb7, parece ser extremamente similar à linhagem 9a5c, uma vez que as diferenças genômicas detectadas entre estas duas linhagens são restritas a um pequeno número de ORFs ausentes, distribuídas ao longo do cromossomo principal e do plasmídeo pXF51 (da Silva *et al.*, 2007).

As principais diferenças entre os genomas de *X.fastidiosa* são regiões derivadas de bacteriófagos. Estas regiões são responsáveis por rearranjos cromossomais e deleções em linhagens de *X. fastidiosa*, em pelo menos 16 regiões cromossomais nos genomas que estão translocadas e/ou invertidas. Sequências relacionadas a fagos compreendem entre 9% e 15% dos genomas de *Xylella* e ~ 47 elementos de fagos foram identificados nos genomas. A maioria dos elementos de prófagos parece ser um resultado de eventos independentes de inserção, implicando que fagos ativos são frequentemente encontrados. Além de serem responsáveis por alterações em larga escala na estrutura e organização de genomas, os prófagos também são capazes de carregar genes com funções não diretamente relacionadas com a propagação de fagos e, dessa forma desempenhando um papel decisivo na evolução do genoma deste fitopatógeno (de Mello Varani *et al.*, 2008; Summer *et al.*, 2010).

Regiões compreendendo mais de 10.000 bases e pelo menos 80% de genes relacionados a fagos em sua constituição foram definidas como regiões similares a prófagos; as menores foram definidas como remanescentes de prófagos. A soma de todas as regiões elegidas representa 1.300.341 bp do esqueleto do cromossomo de

X. fastidiosa como sendo de origem de bacteriófago. A linhagem de citros 9a5c contém seis regiões similares a prófagos e cinco regiões remanescentes de fagos somando 342.510 pb (12.53%); a linhagem de videira Temecula1 contém oito regiões similares a prófagos e duas pequenas regiões remanescentes de fagos somando 239.751 pb (9.50%); a linhagem de espiiradeira Ann1 contém dez regiões similares a prófagos e apenas uma região remanescente de fago somando 339.583 pb (12.97%); e a linhagem de amendoeira Dixon contém onze regiões similares a prófagos e com três regiões remanescentes de fagos somando 378.497 pb (14.39%). O comprimento relativo é de um total 32 kpb para as regiões similares a prófago e 7.5 kpb para as regiões remanescentes de fagos (de Mello Varani *et al.*, 2008).

É interessante notar que os genes para empacotamento de DNA e para a cabeça do fago lembram em organização e sequência de fagos similares ao lambda, e os genes para a placa de base, cauda e fibras da cauda (gpV, gpW, gpJ, gpI, gpU, gpX e gpD) lembram fagos similares ao P2 (de Mello Varani *et al.*, 2008).

Os casos mais interessantes de ORFs de prófagos e possivelmente relacionadas a patogenicidade bacteriana, são a proteína “*Zonular occludens toxin-like*” (*zot*) e a proteína E associada a virulência (*vapE*). Os produtos destas ORFs podem estar relacionados com interações entre a planta e a bactéria. Um grupo de proteínas de sistemas toxina/anti-toxina em regiões de prófagos (*higA/higB* e *relE/relB*) também foram encontrados. Estas proteínas são muito comuns em plasmídeos, onde aumentam a estabilidade dos mesmos, e em cromossomos bacterianos (provavelmente por aquisição de fagos ou incorporação de plasmídeos), onde elas estão relacionadas com a estabilização de genomas de fagos no cromossomo do hospedeiro. Neste caso, o fago pode interferir com a transcrição e tradução de genes do hospedeiro (de Mello Varani *et al.*, 2008).

Devido ao seu caráter fastidioso, muitos aspectos biológicos da bactéria incluindo a ocorrência de fagos são difíceis de analisar. Análises de sequências de genomas inteiros de linhagens de *X. fastidiosa* identificaram muitas sequências de prófagos, incluindo fagos putativos classificados como *Siphoviridae*, *Podoviridae* e *Inoviridae* (Chen e Civerolo, 2008). O mais importante e mais estudado fago dentro da família *Siphoviridae* é o fago lambda, amplamente encontrado nos cromossomos de enterobactéria, onde ele desenvolve diversos papéis biológicos, sendo a maioria deles relacionados à aquisição de genes de virulência pela bactéria através de transferência gênica lateral (de Mello Varani *et al.*, 2008).

Chen e Civerolo (2008) observaram partículas icosaédricas fora das células e aderidas a elas em *X. fastidiosa*. Caudas bem definidas não foram aparentes,

embora uma estrutura muito pequena e fina lembrando uma cauda similar a fago foi observada. A morfologia isométrica e o tamanho destas partículas sugerem que elas sejam vírions putativos de bacteriófagos, provavelmente da família *Podoviridae*. Tais partículas icosaédricas sem cauda podem ser agrupadas em dois tipos: do tipo grande com cerca de 45 nm, com nenhuma cauda pequena observada e cristas na superfície das partículas, provavelmente pertencendo a *Podoviridae*; e do tipo pequena com cerca de 30 nm, que algumas vezes foram vistas em uma cadeia ordenada, provavelmente pertencendo a *Microviridae*. As partículas com cauda poderiam ser *Siphoviridae* e as partículas filamentosas poderiam ser *Inoviridae*. Interessantemente, todas as quatro famílias de fagos foram preditas nos genomas de *X. fastidiosa* baseado em análises de sequência de prófagos (Chen e Civerolo, 2008).

Os fagos em *X. fastidiosa* foram relatados em culturas na fase estacionária tardia ou senescentes. Isto foi baseado na hipótese de que crescimento prolongado em cultura criaria um estresse físico e/ou químico para facilitar a indução de fagos lisogênicos em um ciclo lítico de modo que as partículas de fagos se tornassem visíveis (Chen e Civerolo, 2008).

O conteúdo de fagos pode variar entre linhagens que infectam o mesmo hospedeiro. Chen e colaboradores (2005) relataram uma sequência de DNA de fago de 547 bp do genoma de uma linhagem de videira isolada na Flórida. A sequência compartilhou alta similaridade ao gene de uma integrase da família de fagos *Podoviridae*. Interessantemente, esta sequência está ausente na sequência completa do genoma da linhagem da Califórnia, Temecula1, mas presente em outras linhagens de videira da Califórnia (Chen *et al.*, 2005).

Xylella carrega um dos maiores conjuntos gênicos flexíveis descritos em bactérias, e a análise detalhada da composição deste conjunto certamente será importante para explicar sua eficiente adaptabilidade em infectar tão ampla variedade de espécies vegetais (Nunes *et al.*, 2003; Van Sluys *et al.*, 2003; Chen *et al.*, 2005; Moreira *et al.*, 2005; da Silva *et al.*, 2007; de Mello Varani *et al.*, 2008).

1.3. Fatores de virulência e patogenicidade em *X. fastidiosa*

Os fatores de virulência que *X. fastidiosa* aparentemente utiliza para causar doenças diferem daqueles de outros patógenos vasculares que interagem diretamente com tecidos vegetais vivos. Presumivelmente é a diferença da habilidade do patógeno em se espalhar amplamente dentro de hospedeiros vegetais susceptíveis que os distingue de plantas nas quais a disseminação é mínima e nas

quais o patógeno existe como um endofítico relativamente inofensivo. Embora tal dispersão nas plantas poderia claramente beneficiar interações comensais de endofíticos com a planta, apenas quando a colonização se tornar excessiva em espécies de plantas que tal interação se tornaria patogênica (Chatterjee *et al.*, 2008). Para um patógeno realizar uma interação compatível e bem sucedida com o hospedeiro, ele deve ultrapassar barreiras constitutivas, evitar ou suprimir as respostas imunes, e reprogramar a célula do hospedeiro para o estabelecimento de estruturas de infecção durante penetração, invasão e reprodução (Hok *et al.*, 2010).

Algumas das características de *X. fastidiosa* associadas a sua patogenicidade são o crescimento agregativo e a formação de biofilme (Koide *et al.*, 2004; Chatterjee *et al.*, 2008; Caserta *et al.*, 2010; Voegel *et al.*, 2010). Estas são características comuns no mundo microbiano, que conferem maior resistência a compostos antimicrobianos, aumentam a capacidade das células em adquirir nutrientes do ambiente, tornam mais eficiente a detoxificação devido a um aumento na expressão de bombas de efluxo, além de contribuir com a adesão ao substrato. A formação do biofilme parece ser um processo sequencial no qual células planctônicas aderem às superfícies do hospedeiro, se auto-agregam a outras células e formam uma matriz de biofilme, a qual consiste de exopolissacarídeos, proteínas e provavelmente DNA (Rudrappa *et al.*, 2008; Flemming e Wingender, 2010). Em *X. fastidiosa*, adesinas fimbriais e não-fimbriais são relevantes no desenvolvimento do biofilme e proporcionam a adesão da bactéria aos tecidos do hospedeiro e do inseto vetor (Chatterjee *et al.*, 2008; Caserta *et al.*, 2010; Voegel *et al.*, 2010).

O operon *gum* em *Xylella*, homólogo ao das espécies de *Xanthomonas*, é responsável pela produção e secreção de exopolissacarídeo. Este produto é responsável por formar a base da estrutura do biofilme bacteriano em seus hospedeiros vegetais (Lambais *et al.*, 2000). O genoma de *Xylella* contém genes altamente relacionados com o operon *gumBCDEFHJKM*. *GumD*, *GumH*, *GumK* e *GumM* são conhecidos por estarem envolvidos na montagem do intermediário pentassacarídico lipídico. *GumB*, *GumC*, *GumE* e *GumJ* especula-se que tenham papel na polimerização e translocação do exopolissacarídeo. Entretanto, *gumI* (transferase que incorpora a manose terminal), *gumL* (transferase de cetil piruvato) e *gumG* (acetil transferases) estão ausentes, sugerindo *X. fastidiosa* produza um polissacarídeo modificado (Dow e Daniels, 2000; Chatterjee *et al.*, 2008).

O genoma de *X. fastidiosa* também contém, ainda que parcialmente, o *cluster* de genes *rpf*, responsável pela síntese, secreção e regulação de DSF, uma molécula sinalizadora. Em *Xanthomonas*, este *cluster* compreende nove genes (*rpfA-I*). *rpfA*

codifica uma aconitase implicada em homeostase de ferro; *rpfB* codifica uma ligase de acil-CoA de cadeia longa, e junto com *rpfF*, que codifica uma enoil CoA-hidratase, estão envolvidos na síntese e processamento de DSF. *rpfC* e *rpfG* codificam membros de um sistema de dois componentes, sendo RpfC, o sensor (histina quinase e *phospho-relay*) e RpfG, o regulador da resposta, que possui um domínio HD-GYP (atividade de fosfodiesterase de c-di-GMP). *rpfH* é estruturalmente relacionado com o domínio sensorial codificado por *rpfC* (Chatterjee *et al.*, 2008; Seshasayee *et al.*, 2010; Ryan e Dow, 2011). Em *X. fastidiosa*, DSF modula negativamente os níveis de c-di-GMP, resultando na agregação celular e formação de biofilme (Chatterjee *et al.*, 2008) ao contrário do que se observa em outras bactérias (Seshasayee *et al.*, 2010; Ryan e Dow, 2011).

Para que consiga se espalhar pelos vasos do xilema e colonizar a planta de forma sistêmica, *X. fastidiosa* utiliza a motilidade *twitching* dependente de *pilus* do tipo IV. Seu genoma codifica genes relacionados para biogênese desta classe de *pilus* incluindo a adesina fimbrial localizada na extremidade do pilus (Meng *et al.*, 2005). Motilidade microbiana pode ser controlada por quimiotaxia, um processo no qual os microorganismos sentem o estímulo no ambiente e se movem em direção ou para longe do sinal ambiental. Ortólogos dos genes *pilLGIJ* e *chpBC* foram identificados no genoma de *X. fastidiosa* e aparentemente codificam um sistema quimiossensorial envolvido no controle direcional e da taxa de “*twitching motility*” em resposta a variação no ambiente ao seu redor, chamado de operon Pil-Chp. Este operon também tem um efeito positivo na formação do biofilme (Cursino *et al.*, 2011).

A disseminação sistêmica do patógeno é limitada pelas membranas do poro que separam vasos do xilema de seus vizinhos, e talvez pela produção de tiloses e géis ricos em polissacarídeo pelas plantas que bloqueiam as células do xilema após a infecção. *X. fastidiosa* é uma bactéria com dimensões variando entre 250 a 500 e 1.000 a 4.000 nm, e portanto, grandes para atravessar livremente através da maioria dos poros que tenham sido descritos em angiospermas (variam de 5 a 20 nm). *X. fastidiosa* secreta enzimas que degradam a membrana do poro, como por exemplo, endoglucanases (EGase) e poligalacturonases (PG). A ação combinada destas duas enzimas se mostrou necessária para aumentar o diâmetro médio dos poros. Dessa forma, PG deve remover as moléculas de pectina para expor xiloglucanos, os quais são clivados pela EGase, alterando a matriz da membrana do poro e permitindo que células bacterianas atravessem (Pérez-Donoso *et al.*, 2007). Um mutante de poligalacturonase de *X. fastidiosa* se mostrou restrito ao ponto de

inoculação e incapaz de se mover sistemicamente na videira, indicando sua participação no movimento entre vasos (Roper *et al.*, 2007). O sistema de secreção do tipo II parece ser responsável pela secreção dessas enzimas extracelulares. Alternativamente, degradação localizada das paredes celulares da planta pode liberar carboidratos necessários para o crescimento bacteriano no xilema (Lambais *et al.*, 2000; Chatterjee *et al.*, 2008). Por outro lado, vale destacar que cepas de citros e café aparentemente não expressam PG como consequência de uma mutação em sua sequência codificadora (Simpson *et al.*, 2000; Van Sluys *et al.*, 2003).

Durante a infecção por *X. fastidiosa*, elicitores das respostas de defesa da planta sintetizados pela bactéria e/ou liberados pela parcial degradação da parede celular pode induzir a geração de espécies reativas de oxigênio, tais como superóxido e peróxido de hidrogênio. Vários genes envolvidos em detoxificação de espécies reativas de oxigênio têm sido identificados em *X. fastidiosa*, incluindo genes que codificam para catalase, superóxido dismutase, glutathione peroxidase e glutathione S-transferase, assim como genes regulatórios como *oxyR* (Toledo *et al.*, 2011) e *Ohr* (Cussiol *et al.*, 2003). Isto sugere que *X. fastidiosa* pode responder aos estresses oxidativos e contra-atacar uma das mais rápidas resposta da planta à infecção do patógeno.

Entre os genes para toxinas identificados em *X. fastidiosa*, quatro são similares a hemolisinas pertencendo a família RTX, que são toxinas formadoras de poros que interagem com a membrana plasmática do hospedeiro, causando vazamento de compostos celulares (Simpson *et al.*, 2000). Além disso, genes que codificam colicinas do tipo V são expressos em *X. fastidiosa* (de Souza *et al.*, 2003; de Souza *et al.*, 2004; Pashalidis *et al.*, 2005; Zaini *et al.*, 2008) sugerindo que estas desempenhem algum papel na virulência e patogenicidade desta bactéria.

Mudanças fenotípicas são especialmente importantes para patógenos com histórias de vida complexas, os quais devem sobreviver em múltiplos ambientes sob diferentes estresses dependentes do habitat, como é o caso de *X. fastidiosa* (Chatterjee *et al.*, 2008). Por exemplo, mudanças na expressão gênica podem ser induzidas pela presença de carboidratos do hospedeiro e/ou do inseto vetor. Killiny e Almeida (2009) mostraram que dois polissacarídeos estruturais, pectina e glucano, induziram regulons que afetam o fenótipo deste patógeno. Uma vez que a colonização de *X. fastidiosa* em plantas é heterogênea, alguns vasos podem conter células em um estado não-transmissível, enquanto que outros estejam totalmente colonizados e estão com um fenótipo de transmissão ao vetor (Killiny e Almeida,

2009). Similar aos vasos do xilema nas plantas, a superfície do tubo intestinal do vetor também contém polissacarídeos estruturais. O exoesqueleto dos insetos tem uma composição química complexa, mas é primariamente uma matriz de quitina e proteínas cobertas por uma epicutícula fina que inclui mucopolissacarídeos. Evidências foram apresentadas apoiando a degradação por *X. fastidiosa* e o uso de quitina como fonte de carbono. A presença de quitina no ambiente resultou em uma maior formação de biofilme, sugerindo que este polissacarídeo induz mudanças na regulação gênica de *X. fastidiosa* (Killiny *et al.*, 2010).

No interior das plantas, *X. fastidiosa* está exposta a uma variedade de fatores de estresse, e para assegurar a sobrevivência, *X. fastidiosa* responde a estas situações de estresse através de mecanismos regulatórios específicos. Entre os fatores transcricionais globais pode-se citar o fator sigma 54 (σ^{54}), o qual foi demonstrado atuar na regulação de genes envolvidos na biogênese dos pilus do tipo I e do tipo IV, além de estar envolvido na formação de biofilme em *X. fastidiosa* (da Silva-Neto *et al.*, 2008). Outros dois genes que codificam proteínas com funções regulatórias foram descritos recentemente: *algU* e *gacA*. Um mutante do gene *algU* foi construído e mostrou uma redução nos fenótipos de agregação celular, adesão e formação de biofilme, além de mostrar menor virulência em videiras (Shi *et al.*, 2007). O sistema regulatório de dois-componentes GacS/GacA regula genes reguladores de fatores de patogenicidade e genes envolvidos em “quorum sensing”, produção de toxina, motilidade, formação de biofilme, e produção de polissacarídeo extracelular (Shi *et al.*, 2009). O próprio ambiente do xilema pode fornecer sinais que alteram a expressão de genes e consequentemente o fenótipo. Zaini e colaboradores (2009) mostraram que a seiva coletada no começo da primavera suporta uma cultura estável de *X. fastidiosa* e influencia fortemente na adesão e formação de biofilme. A fase estacionária foi alcançada mais rapidamente no cultivo em 100% de seiva, possivelmente devido à limitação de nutriente. Além disso, diferenças nas taxas de crescimento, maior adesão ao substrato e biofilme mais denso foram observados na presença de seiva (Zaini *et al.*, 2009). Outro estudo com um meio de cultura similar à composição da seiva xilemática mostrou novamente uma expressão aumentada de genes diretamente associados com patogenicidade, virulência e adaptação (Ciraulo *et al.*, 2010).

Outra característica recentemente descrita que pode ser importante na virulência da bactéria seria a produção e secreção de vesículas de membrana externa. Estas vesículas talvez sejam responsáveis pelo transporte de fatores de virulência (toxinas, proteases, adesinas) a localidades distantes e poderiam ser

responsável pelo surgimento de sintomas em folhas não colonizadas por *X. fastidiosa*. Essas vesículas membranares são menores que a bactéria e poderiam, portanto, passar mais facilmente através das membranas do poro e potencialmente viajar longas distâncias no xilema dos hospedeiros vegetais (Voegel *et al.*, 2010).

1.4. Estratégias de controle das doenças causadas por *Xylella*

Um método eficiente e muito usado na prevenção de doenças causadas por *X. fastidiosa* tem sido a busca de variedades naturais e/ou geneticamente selecionadas de plantas que sejam mais resistentes a infecção. No caso de videiras, quase todos os cultivares populares do tipo européia (*V. vinifera*), do tipo americana (*V. labrusca*) e híbrido francesa-americana são susceptíveis à Doença de Pierce. Porém entre cultivares parece existir um gradiente de tolerância a doença, variando do mais resistente até o mais susceptível. A resistência a doença foi encontrada em espécies de *Vitis* nativas do sudeste dos Estados Unidos (Hopkins e Purcell, 2002). Análises comparativas revelaram diferenças significativas na expressão gênica de tecidos de caules de genótipos susceptíveis e resistentes à Doença de Pierce e proteínas potencialmente relacionadas à resistência e susceptibilidade foram identificadas (Yang *et al.*, 2011).

Outros esforços têm sido direcionados no sentido de controle das populações de insetos vetores. A grande variedade de espécies capazes de transmitir a bactéria dificulta ao máximo o seu controle, uma vez que diferentes insetos apresentam diferentes susceptibilidades contra inseticidas usados nos pomares. Além disso, taxas de visitas às árvores de diferentes espécies também podem influenciar na dosagem e frequência da aplicação dos inseticidas. Muitas espécies de gramíneas e plantas ornamentais e/ou de florestas, sem interesse comercial direto, podem ser visitadas pelos insetos e abrigar fontes de inóculos dessas bactérias, dificultando ainda mais o controle dessas populações. Existe um potencial para controle das populações de insetos vetores pela manipulação das condições que influenciam a química do fluido xilemático, gerando plantas inadequadas para alimentação. Enquanto que adultos se alimentam de espécies de hospedeiros diferentes, ninfas jovens sobrevivem e se desenvolvem em apenas algumas poucas espécies de plantas, o que pode limitar também o desenvolvimento destes vetores até a fase adulta. Além das densidades populacionais, a atividade do vetor, dispersão temporal e espacial dos insetos, e seleção de locais nos hospedeiros para a alimentação parecem ser mais importantes para a epidemia da doença do que a eficiência na transmissão e devem ser consideradas em qualquer programa de

controle de pragas e doenças agrícolas (Redak *et al.*, 2004; Janse e Obradovic, 2010).

Algumas práticas de manejo são importantes para manutenção da planta saudável. Em muitos hospedeiros, sintomas causados por *X. fastidiosa* se desenvolvem com maior severidade em plantas enfraquecidas ou mal-cuidadas. Fatores de estresse que podem favorecer o desenvolvimento de doenças incluem seca, poda da raiz com equipamento de cultivo, outras doenças associadas, superprodução de frutos e senescência normal da planta (Hopkins e Purcell, 2002). Outras práticas que se mostram eficientes são as inspeções regulares dos pomares, podas após remoção de árvores doentes, controle de ervas-daninhas dentro e ao redor dos pomares e a eliminação de árvores lenhosas, especialmente carvalhos perto dos pomares (Janse e Obradovic, 2010).

Além disso, diferenças ambientais, como temperatura, água e características do solo também podem afetar a susceptibilidade à doença. Dessa forma, exposição a temperaturas mais baixas pode aumentar a possibilidade de cura. Quando plantas são submetidas a um estresse de baixa temperatura, a fluidez da membrana diminui, o metabolismo é diminuído, e o estresse oxidativo se inicia. Quando a temperatura cai abaixo do congelamento, o estresse para a planta é devido à desidratação, solutos celulares são concentrados, potenciais de membrana mudam, e gelo extracelular pode ocorrer. Todos esses fatores, entre outras mudanças fisiológicas que ocorrem nas plantas em resposta a temperaturas frias, podem afetar a sobrevivência de *X. fastidiosa* no tecido do xilema. Além disso, compostos fenólicos são produzidos pelas plantas em resposta a fatores bióticos e abióticos, incluindo temperaturas baixas, sendo alguns fenólicos podem apresentar propriedades anti-microbianas (Meyer, 2010).

Uma prática cada vez mais utilizada é o uso de uma linhagem menos virulenta ou avirulenta para o controle da população de uma linhagem virulenta. Linhagens fracamente virulentas se multiplicam e se movem sistemicamente, embora mais lentamente na planta, produzindo apenas sintomas menores no hospedeiro. Linhagens de *X. fastidiosa* menos virulentas isoladas naturalmente poderiam ser usadas para desenvolver um sistema de controle biológico. Alternativamente, genes específicos que estão envolvidos na patogenicidade de *X. fastidiosa* poderiam ser mutados para obtenção de linhagens atenuadas. (Hopkins e Purcell, 2002). Um estudo recente mostra a utilização de uma linhagem biocontrole, EB92-1, isolada de sabugueiro (Hopkins *et al.*, 2010) e cujo genoma incompleto foi recentemente elucidado (Zhang *et al.*, 2011). Esta cepa é capaz de colonizar

videiras, mas sem causar sintomas. Um pré-tratamento da planta com essa linhagem gera diferenças significativas entre as vinhas tratadas e as não-tratadas no que diz respeito aos sintomas gerados. Sintomas pareceram ser mais severos em vinhas não-tratadas do que em vinhas tratadas com EB92-1. Entretanto, EB92-1 não influenciou a incidência da doença de Pierce durante o primeiro ano, mas forneceu bom controle nos segundos e terceiros anos (Hopkins *et al.*, 2010).

Outro fator que pode influenciar a resistência às doenças é a natureza da comunidade microbiana endofítica colonizando plantas individuais. Bactérias endofíticas são aquelas que vivem nas partes internas das plantas sem causar aparentes danos aos seus hospedeiros. Araújo e colaboradores (2002) sugeriram que pode existir uma interação entre *Curtobacterium flaccumfaciens*, espécie de *Methylobacterium* e *X. fastidiosa* (Araújo *et al.*, 2002). Crescimento de alguns isolados de *Xylella fastidiosa* em PW pareceu ser estimulado por isolados de *Methylobacterium*. Inibição de *X. fastidiosa* foi observada na presença do sobrenadante de *C. flaccumfaciens*, sendo que altas concentrações resultaram em crescimento reduzido de *X. fastidiosa* (Lacava *et al.*, 2004).

Uma vez que a ocorrência de *X. fastidiosa* é restrita ao continente americano, a alternativa de controle para os demais continentes é implementar procedimentos de quarentena e fitossanitários de modo a prevenir a introdução do patógeno. Considerando a ampla abrangência de hospedeiro, inespecificidade quanto aos insetos vetores e o movimento global do material vegetal, estas medidas administrativas e preventivas devem ser fortificadas por outras ações profiláticas (Janse e Obradovic, 2010).

O desenvolvimento de estratégias para o controle das doenças causadas por *X. fastidiosa* tem sido motivado pela importância econômica das culturas que podem ser afetadas por este patógeno. Segundo dados do Programa de Controle da Doença de Pierce da Califórnia, esta doença ameaça negócios diretamente derivados da vinicultura estimados em mais de 33 bilhões de dólares (<http://www.cdfa.ca.gov/gwss/>). No Brasil, a citricultura é um dos setores mais competitivos do agronegócio. O Brasil produz mais de 30% da laranja cultivada no mundo e 58% do suco de laranja concentrado congelado consumido no mundo, sendo que 97% de todo o suco de laranja produzido no Brasil é exportado. A cadeia produtiva da laranja gera mais de 200 mil empregos diretos em mais de 300 municípios e uma receita cambial de exportação que varia entre US\$ 1,5 bilhão e US\$ 2,5 bilhões anuais (Neves *et al.*, 2007; Neves *et al.*, 2011). Segundo levantamentos realizados pelo Fundo de Defesa da Citricultura, a CVC atingia, em

2005, 43% das plantas do parque citrícola do Estado de São Paulo, passando a 39% em 2009. Esta queda da incidência da doença deveu-se possivelmente à produção de mudas sadias, aumento do uso de inseticidas para controle do inseto vetor e a eliminação de pomares afetados pelo amarelinho. Entretanto, dados recentes do Fundecitrus indicam aumento da incidência e da severidade desta doença, visto que a proporção de árvores doentes no Estado de São Paulo passou de 35,5% em 2010 para 40,3% em 2011. Este aumento é preocupante e revela que muitos citricultores deixaram de adotar as medidas regulares de prevenção (<http://www.fundecitrus.com.br>).

2. Objetivos

Como discutido anteriormente, genomas de cepas de *X. fastidiosa* isoladas de distintos hospedeiros já foram sequenciados completa ou parcialmente: 9a5c de citros; Temecula-1 e GB514 de videira; Dixon, M12 e M23 de amendoeira; Ann-1 de espiroleira e EB92-1, isolada de sabugueiro do Canadá e utilizada como bio-controle para Doença de Pierce de videiras (Simpson *et al.*, 2000; Bhattacharyya *et al.*, 2002; Van Sluys *et al.*, 2003; Chen *et al.*, 2005; Schreiber *et al.*, 2010; Zhang *et al.*, 2011). A comparação destes genomas entre si e com genomas de outros fitopatógenos, associada a abordagens de genômica funcional e de genética molecular, têm possibilitado o estudo detalhado de mecanismos e fatores potencialmente relevantes na colonização de plantas e insetos por este fitopatógeno e no desenvolvimento de sintomas associados a doenças específicas em seus respectivos hospedeiros vegetais.

Exceto o genoma de 9a5c, todos os demais genomas conhecidos são de cepas isoladas na América do Norte. Assim, com o propósito de contribuirmos para a identificação de genes potencialmente associados à adaptação ao hospedeiro e a virulência de *X. fastidiosa*, nosso grupo propôs a realização do sequenciamento do genoma de cepas isoladas na América do Sul, sendo três delas isoladas de citros: J1a12 (não-virulenta), U24d (virulenta, representante de um haplótipo atualmente predominante no Estado de São Paulo), Fb7 (virulenta, isolada na Argentina); uma cepa isolada de cafeeiro (3124); uma cepa isolada de ameixeira (Pr8x) e outra isolada de hibisco (Hib4). Neste contexto, tivemos neste projeto de Mestrado os seguintes objetivos específicos:

1. Concluir o sequenciamento do genoma da cepa J1a12 que exibe fenótipo não-virulento em citros, o qual havia sido iniciado em nosso grupo (da Silva, 2010);
2. Validar experimentalmente a montagem proposta para o genoma da cepa J1a12;
3. Realizar o sequenciamento do genoma das cepas Pr8x e Hib4, isoladas, respectivamente, de ameixeira e hibisco;
4. Realizar análises comparativas das sequências destes genomas com genoma da cepa 9a5c;
5. Associar diferenças genotípicas das cepas J1a12, Pr8x e Hib4 aos seus respectivos fenótipos.

3. Procedimentos experimentais

3.1. Manutenção e cultivo de *Xylella fastidiosa*

A cepa 9a5c (Li *et al.*, 1999; Simpson *et al.*, 2000) e J1a12 (Monteiro *et al.*, 2001; Koide *et al.*, 2004) de *X. fastidiosa* foram cedidas pelo Fundecitrus (Araraquara, SP). O cultivo destas cepas foi realizado em meio complexo PWG que é uma modificação do meio PW (Davis, 1981) e é constituído de fitona peptona 4g/L; peptona de caseína com digestão triptica 1g/L; cloreto de hemina 0,001%; K₂HPO₄ 1,2g/L; KH₂PO₄ 1g/L; MgSO₄.7H₂O 0,4g/L; glutamina 0,4% e glicose 0,5%. Para o cultivo em meio líquido, as bactérias, após sete dias de cultivo em meio sólido PWG-ágar (PWG contendo ágar 1,5%), foram transferidas para 50 mL de meio PWG e mantidas a 28°C com 100 rpm de agitação por mais sete dias.

A cepa Pr8x foi isolada de ameixeira, espécie *Prunus salicina*, no município de Jarinu no estado de São Paulo pela aluna de doutorado Cristiane Muller do Departamento de Entomologia da Escola Superior de Agricultura “Luis de Queiroz”, USP, Piracicaba, SP. A cepa Hib4 de hibisco (*Hibiscus rosa-sinensis*) foi isolada no Centro de Citricultura Sylvio Moreira, Cordeirópolis, SP a partir de material vegetal coletado em Campinas, SP, pelo Prof. Dr. Elliot Kitajima (ESALQ/USP). O cultivo destas duas cepas também foi realizado em meio PWG ou em meio BCYE (Wells *et al.*, 1981) nas condições descritas acima. O meio BCYE é constituído de extrato de levedura 10g/L; carvão ativado 2g/L; tampão ACES 10g/L; L-cisteína HCL 0,4g/L; pirofosfato férrico 0,25g/L e ágar 17g/L.

Estoques de células são mantidos em PWG contendo 50% de glicerol em freezer a -80°C. A manutenção das culturas é feita em placas de meio PWG-ágar com repiques semanais e incubação a 28°C, sendo o número de passagens controlado. Após 30 passagens, novo cultivo é iniciado a partir de estoques congelados a -80°C.

3.2. Avaliação morfológica de colônias de *X. fastidiosa*

Para a avaliação morfológica das colônias, as diferentes cepas de *X. fastidiosa* foram cultivadas em meio PWG líquido por sete dias sob agitação branda (80-100 rpm) e temperatura de 28°C. Após este período, 100 µL da suspensão foram coletados e homogeneizados, seguindo-se deposição em placas de PWG-ágar por gotejamento. A morfologia das colônias foi examinada após 4-5 dias de crescimento a 28°C. Para tal porções do ágar foram cortadas e transferidas para lâminas de microscópio. As lâminas foram observadas em um microscópio óptico

invertido de transmissão “*Axiocvert 200*” (*Carl Zeiss MicroImaging*) em um aumento de 20x. As imagens foram capturadas por uma câmera “*Canon*” e registradas com o *software* “*Scopeimage 9.0*”.

3.3. Isolamento de DNA total

Para purificação de DNA total bacteriano foi utilizado o “*Wizard Genomic DNA Kit*” (Promega) seguindo as instruções do fabricante com pequenas modificações. Usualmente partimos de 50 mL de cultura de *X. fastidiosa* obtida em PWG, os quais foram centrifugados a 13.000 xg por 2 min e o precipitado obtido foi ressuspensão em ~3,5-4mL de tampão de lise do kit. A suspensão foi, então, incubada a 80°C por 5 min, seguindo-se incubação em temperatura ambiente por 5 min. Aliquotas de 600 µL do lisado foram transferidas para tubos de 1,5 mL, seguindo-se adição 3 µL da solução de RNase, inversão do tubo algumas vezes e incubação a 37°C por 60 min. O tubo é transferido para temperatura ambiente por 5 min e em seguida, foram adicionados 200 µL da solução de precipitação de proteína e o tubo foi agitado vigorosamente por 20 seg. Após incubação em gelo por mais de 5 min, a amostra foi então submetida à centrifugação a 16.000 xg por 3 min. O sobrenadante contendo o DNA foi transferido para um tubo de 1,5 mL contendo 600 µL de isopropanol à temperatura ambiente. A amostra foi misturada gentilmente até visualização da massa de DNA. Em seguida, a solução contendo o DNA foi submetida à centrifugação de 16.000 xg por 2 min. O sobrenadante foi removido e o líquido residual foi drenado em papel absorvente. Posteriormente, ao tubo contendo o precipitado de DNA, foram adicionados 600 µL de etanol 70% à temperatura ambiente para lavagem do precipitado. Novamente a amostra foi centrifugada em 16.000 xg por 2 min e o etanol foi aspirado completamente. O líquido residual foi drenado em papel absorvente e o precipitado seco ao ar em câmara de fluxo laminar. Por fim, foram adicionados 100 µL de solução de hidratação de DNA (Tris.HCl 10mM, pH7,5; EDTA 1mM), seguindo-se incubação a 65°C por mais de 1h para completa solubilização do DNA. O DNA total obtido foi armazenado a 4°C.

A quantificação e avaliação da qualidade do DNA foi feita através da determinação das absorbâncias (Abs) a 260nm e 280nm em espectrofotômetro ND-1000 (*NanoDrop Technologies*), sendo que Abs_{260nm} = 1 equivale a 50 µg de DNA fita dupla/mL.

As amostras de DNA extraídas de *X. fastidiosa* foram analisadas quanto a sua integridade através de eletroforese convencional em géis de agarose conforme descrito no item 3.6.1.

3.4. Isolamento de DNA plasmidial

Para a purificação do DNA plasmidial foi utilizado o “*PureLink HiPure Plasmid DNA Purification kit*” (Invitrogen) segundo as instruções do fabricante. Antes de iniciar o procedimento foi adicionada uma alíquota de RNase A ao tampão de ressuspensão e a coluna *HiPure Midi* foi equilibrada com a adição de 10 mL de tampão de equilíbrio à coluna. Em seguida, o precipitado de células foi obtido pela centrifugação a 4000 rpm por 5 min de uma cultura de *X. fastidiosa* em meio PWG líquido de 50 mL. O precipitado de células foi ressuspenso em 4 mL do tampão de ressuspensão (contendo RNase) e em seguida foram adicionados 4 mL do tampão de lise. A suspensão foi misturada gentilmente invertendo-se o tubo e o lisado foi incubado à temperatura ambiente por 5 min. Após este período, foram adicionados 4 mL de tampão de precipitação, misturando-se delicadamente a suspensão até torná-la homogênea. A amostra foi então centrifugada por 10 min a 15.000 xg, à temperatura ambiente, o sobrenadante foi removido cuidadosamente e adicionado à coluna “*HiPure Midi*” (previamente equilibrada). A coluna foi lavada duas vezes com 10 mL de tampão de lavagem e o DNA plasmidial foi eluído com 5 mL de tampão de eluição. Para a precipitação do DNA, foram adicionados 3,5 mL de isopropanol ao eluato, seguindo-se centrifugação por 30 min a 15.000 xg a 4°C. O sobrenadante foi descartado com cuidado e foram adicionados 3 mL de etanol 70% ao precipitado. A amostra foi novamente centrifugada a 15.000 xg a 4°C por 5 min. O sobrenadante foi descartado e o precipitado foi seco ao ar em câmara de fluxo laminar por 10 min ou até eliminação completa do álcool. Por fim, o precipitado foi ressuspenso em 100-200 µL de tampão TE (Tris.HCl 10mM, pH7,5; EDTA 1mM) e armazenado a -20°C. A quantificação e avaliação da qualidade do DNA foram feitas como descrito no item 3.3.

3.5. Reação em cadeia da polimerase (PCR)

3.5.1. Genotipagem das cepas

A genotipagem das cepas de *X. fastidiosa* foi realizada através de ensaios de PCR com os pares de *primers* listados na Tabela 1. Foram utilizados *primers* para as regiões externas às CDS XF0077, que codifica um precursor de adesina de fimbria (produtos de 1599 pb em 9a5c e 549 pb em J1a12) e XF1250, que codifica uma

arginina desaminase (produtos de 1122 pb em 9a5c e 159 pb em J1a12) e que detectam diferenças genômicas previamente caracterizadas entre as cepas 9a5c e J1a12 (Koide *et al.*, 2004). O par de *primers* para amplificar uma região da CDS XF1952 (Tabela 1) foi utilizado para a genotipagem da cepa Hib4 e foi planejado a partir da sequência de seu genoma descrita neste trabalho. O par de *primers* CVC-1 e 272-2-int é adequado para genotipagem de cepas *X. fastidiosa* de citros e de café (Pooler e Hartung, 1995) e gera produtos de 500 pb na PCR.

Tabela 1. Sequências dos pares de *primers* para genotipagem de *X. fastidiosa*

Primer	Direto (F)	Reverso (R)
XF0077	CTAAGCGCGACAATGAGGA	GTGGTGTGTTGGGCAAGCTAT
XF1250	ATGGCAAGGTGGAAATCAG	CAACGCACATCCCCTAAGTC
XF1952	GTGGCTACAGAATCTATGTCA	TCCGGCATCTCAGTCATAGC
CVC-1/ 272-2-int	AGATGAAAACAATCATGCAAA	GCCGCTTCGGAGAGCATTCTCT

Para os ensaios de PCR, 100ng de DNA isolado das cepas de interesse foram misturados com 1 µL de solução de dNTPs 10 mM (Prodinol), 1,5 µL de MgCl₂ 25 mM, 10 µL de tampão 10x da Taq DNA Polimerase (Invitrogen), 1 µL de cada *primer* na concentração de 10 pmol/µL e água MilliQ estéril para o volume final da reação de 50 µL. As reações com os pares de *primers* para as CDSs XF0077 e XF1250 foram submetidas à desnaturação a 95°C, 5 min e a 30 ciclos de 95°C, por 45 seg; 60°C, por 30 seg; 72°C, por 2 min. As reações com o par de *primers* da CDS XF1952 e com o par de *primer* CVC-1/272-2-int foram submetidas à desnaturação a 95°C, 5 min e a 30 ciclos de 95°C, por 45 seg; 62°C, por 30 seg; 72°C, por 2 min. Ao final destes ciclos, as reações foram incubadas a 72°C por 10 min e armazenadas a 4°C até sua análise por eletroforese em gel de agarose 1%. O controle da reação de PCR foi realizado substituindo-se a quantidade de DNA molde por água mili-Q. As PCRs foram realizadas em um termociclador “GeneAmp PCR System 9700” (Applied Biosystems) e os produtos obtidos (4 µL da reação) foram avaliados por eletroforese em gel de agarose 1%.

3.5.2. Análise de pXF27 e pXF51

A presença dos plasmídeos pXF27 e pXF51 em cepas de *X. fastidiosa* foi avaliada por PCR utilizando-se os pares de *primers* listados na Tabela 2. Foi utilizado um par de *primers*, pXF27_circ, para as regiões inicial e final do pXF27, de modo a observar se a reação de PCR geraria o produto esperado, de 3000 pb, sugerindo que o plasmídeo estaria circular. A mesma estratégia foi adotada para o

plasmídeo pXF51, utilizando o par de *primers* pXF51_circ para verificar sua circularização, que geraria um produto de 902 pb após o ensaio.

Outros dois pares de *primers*, pXF27_8 e pXF51_vapD foram utilizados para amplificação de região correspondente, respectivamente às CDS VirD4, presente no pXF27 segundo anotação automática do servidor RAST (Aziz *et al.*, 2008) e XFa0052 de pXF51 (Marques *et al.*, 2001). No primeiro caso, produto gerado é 1181 pb e no segundo de 392 pb.

Os ensaios de PCR foram realizados como descrito no item anterior, sendo que com o par de *primers* pXF27_circ, as reações foram submetidas à desnaturação a 95°C, 5 min e a 30 ciclos de 95°C, por 45 seg; 58°C, por 30 seg; 72°C, por 3 min e 30 seg. Com os iniciadores para a CDS VirD4, as reações foram submetidas à desnaturação a 95°C, 5 min e a 30 ciclos de 95°C, por 45 seg; 60°C, por 30 seg; 72°C, por 2 min. Com ambos os pares de *primers* para o pXF51 (pXF51_circ e pXF51_vapD) as reações foram submetidas à desnaturação a 95°C, 5 min e a 30 ciclos de 95°C, por 45 seg; 55,5°C, por 30 seg; 72°C, por 2 min.

Tabela 2. Sequências dos pares de *primers* utilizados para detecção de pXF27 e pXF51 em cepas de *X. fastidiosa*

Primer	direto (F)	reverso (R)
pXF27_circ	ACTCGAAATAATAGCGGTCGTC	GTTCGACCAATACAGCACTACAG
pXF27_8	TCGCACAGACTAGCGACAAG	CGTGTAGTTAAAGCTCCCGAGT
pXF51_circ	ATGATGGCAGCCTTACCAAG	CCATAGTTTGTGTGGCAACC
pXF51_vapD	TCGCTGCCTAATCGTTTTTG	GTTGGCCAATGATTTCTTCG

3.6. Eletroforese de DNA em gel de agarose

3.6.1. Eletroforese convencional

O gel utilizado na eletroforese convencional (Smith, 1991) foi preparado com 1% de agarose em pó diluída em tampão TBE (Tris-base 89 mM, ácido bórico 89 mM, EDTA 2 mM pH 8,0), sendo a mistura aquecida a 100°C para completa dissolução. Antes da solidificação do gel no aparato apropriado, foram adicionados 2,5 µL de solução de brometo de etídio 10 mg/mL para cada 50 mL de solução de agarose. A eletroforese foi realizada em uma cuba eletroforética contendo tampão TBE 1x a 100 V, em direção ao pólo positivo, até a saída do corante indicador. O volume total de amostra aplicado no gel foi de ~10 µL [~100 ng de DNA em tampão de amostra (corante azul de bromofenol 0,25% (p/v); sacarose 40% (p/v); Tris-base 10 mM pH 7,5; EDTA 1 mM)]. Ao final da corrida, o gel foi escaneado em um

fotodocumentador com luz UV e o tamanho dos fragmentos foi estimado com base na comparação à mobilidade de fragmentos de DNA com tamanhos já conhecidos (GeneRuler 1kb DNA Ladder, Fermentas).

3.6.2. Eletroforese longa em baixa voltagem

Esse tipo de eletroforese é semelhante à eletroforese convencional descrita acima, com algumas modificações (Saillard *et al.*, 2008). O gel utilizado nesta eletroforese foi preparado com 0,5% de agarose sendo que foi adicionado solução de brometo de etídio ao tampão de corrida (TBE 1x) na mesma concentração presente no gel. A eletroforese foi realizada a uma voltagem de 20-22V por um período de 45 horas em câmara fria com temperatura constante de ~4-6°C. Se necessário o volume do tampão de corrida foi corrigido para compensar eventual evaporação. Periodicamente, o tampão de corrida foi misturado para igualar as concentrações dos eletrólitos em ambas as extremidades da cuba. A eletroforese em gel de agarose em baixa voltagem foi utilizada como alternativa à eletroforese convencional, com o objetivo de separar mais eficientemente moléculas de DNA de alto peso molecular.

3.7. Digestão do pXF51 e pXF27 com enzimas de restrição

3.7.1. Digestão do produto de PCR pXF27_circ

O produto de 3000 pb obtido com o par de *primers* pXF27_circ a partir de DNA total da cepa J1a12, foi separado por eletroforese convencional, a banda correspondente foi excisada do gel e o DNA foi extraído do bloco de agarose com o “QIAEX II gel extraction kit” (Quiagen). Os blocos de agarose tiveram sua massa aferida e a partir dessa medida, foi adicionado o tampão QX1, na proporção de 3:1 (essa proporção é adequada para fragmentos de DNA entre 100pb e 4000pb). Em seguida, a suspensão da resina “QIAEX II” foi realizada por agitação vigorosa por 30 seg e 10 µL foram transferidos para o tubo, seguindo-se incubação a 50°C por 10 min para solubilização da agarose e ligação do DNA ao “QIAEX II”. A mistura foi submetida à agitação vigorosa a cada 2 min e em seguida centrifugada por 30 seg a 13.000 rpm. O sobrenadante foi cuidadosamente removido com uma pipeta e o precipitado foi lavado com 500 µL do tampão QX1 por agitação vigorosa, seguindo-se centrifugação por 30 seg a 13.000 rpm, sendo qualquer traço de sobrenadante removido. O precipitado foi lavado mais duas vezes com 500 µL de tampão PE (com adição de etanol 100%) com agitação vigorosa e centrifugação. As amostras foram secas ao ar em câmara de fluxo laminar até que o precipitado se tornasse branco. O DNA foi eluído em 20 µL Tris-HCl 10 mM, pH8,5, ou em água e, ressuspendido por

agitação vigorosa. As amostras foram incubadas à temperatura ambiente por 5 min e armazenadas a -20°C.

O fragmento de 3000 pb purificado foi submetido a digestão com a enzima de restrição *SacII* (New England Biolabs), a qual geraria fragmentos 1926 pb e outro de 1074 pb de acordo com predição a partir do mapa de sítios de restrição do pXF27. A reação continha ~130ng de DNA, 2 µL de tampão NEBuffer 4 (acetato de potássio 50 mM, tris-acetato 20 mM, acetato de magnésio 10 mM) (New England Biolabs), 1 µL da solução da enzima (20U/µL) e água para volume final de 20 µL. A reação foi incubada a 37°C por 16 h e interrompida pelo aquecimento a 65°C por 20 min. A digestão do DNA foi analisada em eletroforese convencional em gel de agarose 1% como descrito no item 3.6.1.

3.7.2. Digestão do DNA plasmidial total

A partir da sequência de pXF51 e da sequência proposta para pXF27 construímos mapas de restrição utilizando o *software* gratuito “ApE” (*A plasmid Editor*) (Davis, 2009) e selecionamos as enzimas *NdeI*, *SacII* e *XhoI* para realizar a digestão destes plasmídeos. O DNA plasmidial obtido como descrito no item 3.4 foi submetido a digestões com as enzimas de restrição: *NdeI* (New England Biolabs), *SacII* (New England Biolabs) e *XhoI* (New England Biolabs), em reações contendo ~130ng de DNA, 2 µL de tampão NEBuffer 4 (acetato de potássio 50 mM, tris-acetato 20 mM, acetato de magnésio 10 mM) (New England Biolabs), 1 µL das soluções das enzimas (20U/µL) e água para volume final de 20 µL. As reações foram incubadas a 37°C por 16 h e interrompidas por aquecimento a 65°C por 20 min.

As digestões foram analisadas por eletroforese convencional em gel de agarose 0,8% ou por eletroforese longa em baixa voltagem em gel de agarose 0,5%, como descrito no item 3.6.

3.8. Pirosequenciamento de DNA

A tecnologia de pirosequenciamento de DNA (plataforma 454 Genome Sequencer FLX (454 Life Sciences/Roche) baseia-se no sequenciamento por síntese enzimática, onde a incorporação do nucleotídeo complementar na nova cadeia de DNA que está sendo sintetizada resulta em um sinal luminoso detectável (bioluminescência). Cada um dos nucleotídeos é adicionado isoladamente e de forma alternada em cada ciclo da reação. Caso ocorra a reação de incorporação do nucleotídeo pela DNA polimerase, a partir da sequência do DNA molde, uma molécula de pirofosfato é liberada; esta é utilizada na síntese de ATP pela ATP

sulfurilase, juntamente com o substrato adenosina fosfossulfato (APS). O ATP sintetizado é utilizado na reação da luciferase para produção de luz a partir do substrato luciferina. Na etapa subsequente, a apirase remove os nucleotídeos não incorporados antes do próximo ciclo de sequenciamento (Margulies *et al.*, 2005; Ahmadian *et al.*, 2006; Droege e Hill, 2008).

A metodologia de pirosequenciamento não requer a construção prévia de bibliotecas genômicas, seleção de clones e isolamento de plasmídeos (Margulies *et al.*, 2005; Ahmadian *et al.*, 2006). O DNA genômico total é rompido mecanicamente por nebulização em fragmentos de 500 a 800 pares de base. Aos fragmentos são ligados pequenos adaptadores específicos nas extremidades 3' e 5'. As moléculas de DNA fita simples (ssDNA) devidamente ligadas aos adaptadores são imobilizadas em microesferas, sendo que cada microesfera contém uma única fita de DNA. Esta biblioteca de ssDNA ligado a microesferas é então submetida à emulsificação com reagentes para amplificação por PCR em uma mistura de água e óleo; cada microesfera com uma única molécula de ssDNA ligada, inserida em uma micro-gota aquosa suspensa no óleo, constitui um microreator. Cada fragmento ligado a uma microesfera é então amplificado por PCR. A reação de amplificação de toda a biblioteca de fragmentos é realizada simultaneamente, sendo que os fragmentos amplificados permanecem ligados à mesma microesfera original em cada microreator. Ao final da amplificação cada microesfera está recoberta com cerca de 10 milhões de cópias (clones) do mesmo fragmento de DNA.

Os microreatores são posteriormente quebrados e as microesferas são depositadas e imobilizadas em uma placa (*PicoTiterPlate*, PTP) contendo 1,6 milhões de poços com a dimensão exata para comportar uma única microesfera. A placa é em seguida acoplada ao sequenciador para que as reações de pirosequenciamento ocorram. Após 10 a 20 horas de corrida (dependendo do kit de sequenciamento utilizado), são obtidos dados brutos das imagens dos sinais luminosos gerados a partir da *PicoTiterPlate* e detectados por uma câmera CCD (www.roche-applied-science.com). Após ~10 horas adicionais de análise computacional, as imagens são convertidas em sequências, cada uma derivada das imagens de um poço da placa. O equipamento permite a obtenção de ~1 milhão de sequências (*reads*) com tamanho médio de 400-700 nucleotídeos dependendo da versão do kit de sequenciamento.

O tamanho da biblioteca de DNA genômico e o número de regiões da PTP nas quais será aplicada a biblioteca de DNA dependerão dos objetivos e requisitos do sequenciamento, isto é, total de bases ou total de *reads* desejados. Neste trabalho

nós optamos pelo sequenciamento paralelo de bibliotecas distintas utilizando 2 regiões da PTP, as quais são delimitadas pela colocação de uma máscara (*gasket*).

3.8.1. Pirosequenciamento com a metodologia *Paired End Tags*

O conceito principal da metodologia *Paired End Tags* (PET) é a obtenção de sequências (*tags*) de ambas as extremidades (*paired ends*) de fragmentos de DNA de tamanhos conhecidos (Fullwood *et al.*, 2009). Desta forma as distâncias entre as duas extremidades dos fragmentos podem ser usadas para ordenar *contigs* na montagem da sequência genômica. Além disso, regiões genômicas contendo repetições podem ser orientadas e posicionadas pelas suas conectividades a regiões de sequências específicas oferecidas pelas sequências obtidas pelo sequenciamento com a estratégia PET.

A construção de bibliotecas de *Paired End Tags* envolve a circularização direta de fragmentos de DNA gerados utilizando o nebulizador “*Hydroshear*” (Digilab) que pode gerar fragmentos de tamanhos específicos, como por exemplo, 20.000, 8.000 ou 3.000 pares de base. No caso da cepa J1a12 foram utilizados fragmentos na faixa de 3.000 pb. Às duas extremidades dos fragmentos de DNA são ligados oligonucleotídeos (*linker*) e a circularização dos fragmentos é então realizada com DNA ligase. Em seguida, procede-se fragmentação do DNA por nebulização e as etapas regulares do protocolo de pirosequenciamento na plataforma 454/Roche descrita adiante. As sequências (*reads*) contendo *tag-linker-tag* são identificadas computacionalmente na etapa de montagem dos *contigs*.

O pirosequenciamento com a estratégia PET foi realizado na Unidade de Genômica Computacional Darcy Fontoura no Laboratório Nacional de Computação Científica (LNCC) em Petrópolis, RJ.

3.8.2. Pirosequenciamento com a metodologia *shotgun*

3.8.2.1. Reagentes utilizados

Foram utilizados os reagentes e *kits* da série titânio: GS FLX Titanium Rapid Library Preparation, GS FLX emPCR kit Titanium reagents (SV e LV), GS FLX emPCR kit Titanium emulsion oil (SV e LV), GS FLX emPCR kit Titanium Bead Recovery Reagents, GS FLX Sequencing kit Titanium Buffer CB, GS FLX Sequencing kit Titanium Bead and Wash buffers, GS FLX Sequencing kit Titanium Reagents XLR70 e GS FLX Titanium PicoTiterPlate 70x75 (454 Life Sciences/Roche), além de reagentes e *kits* de outros fabricantes tais como: MinElute PCR Purification Kit (Qiagen);

Agencourt AMPure System (Beckman Coulter); *DNA High Sensitivity LabChip kit* (Agilent) e *Quan-iT PicoGreen* (Invitrogen).

3.8.2.2. Quantificação do DNA genômico

A quantificação das amostras de DNA total purificado das cepas Hib4 e Pr8x destinadas ao pirosequenciamento foi realizada com o reagente “*Quant-it PicoGreen dsDNA Reagent*” (Invitrogen). Para a curva padrão são utilizadas concentrações diferentes de DNA de fago lambda fornecido pelo fabricante (50 ng, 25 ng, 12,5 ng, 6,25 ng, 3,125 ng, 1,563 ng e 0,781 ng). As amostras de DNA das cepas Hib4 e Pr8x foram previamente diluídas para concentrações abaixo de 50ng, com base em medições prévias da quantidade de DNA no espectrofotômetro NanoDrop. Para a quantificação, 100µL da solução do reagente fluorescente “*Quant-it Picogreen dsDNA Reagent*”, previamente diluído 1:200 em TE, são misturados a 100µL das amostras de DNA em uma placa de 96 poços. A placa foi deixada em repouso à temperatura ambiente durante 5 minutos, antes da leitura da fluorescência em um espectrofluorímetro (excitação em 480nm e a intensidade da emissão medida a 520nm). Como esperado, os valores das concentrações de DNA das duas amostras, cepa Hib4 (152 ng/µL) e Pr8x (27 ng/µL) medidas com *PicoGreen* foram inferiores aos valores das concentrações estimadas com o espectrofotômetro.

3.8.2.3. Análise qualitativa do DNA genômico

As amostras de DNA das cepas Hib4 e Pr8x para o pirosequenciamento tiveram sua integridade também analisada pelo “*High sensitivity DNA kit*” (Agilent) utilizando o equipamento *Bioanalyzer* (Agilent). Antes de iniciar a análise das amostras de DNA, os reagentes a serem utilizados foram equilibrados à temperatura ambiente durante 30 minutos. Aliquotas do gel para uma corrida foram filtradas antes da adição de 1,5µL do “*dye*” seguindo-se centrifugação por 10 minutos a 4000 rpm. A solução “*gel-dye mix*”, devidamente protegida da luz, foi utilizada para a montagem do *chip* com o auxílio de um suporte e de uma seringa. Foram adicionados 5µL do marcador que demarca os limites de tamanho de amostras de DNA de cada tipo de kit utilizado no *Bioanalyzer* (no caso do *High Sensitivity*, usado neste ensaio, os limites são 50-7.000 pb). Em seguida, em um dos poços foi adicionado 1µL do *ladder* e nos poços restantes foi adicionado 1µL das amostras de DNA total. O *chip* foi vortexado em seu suporte por 1 minuto a 2400 rpm e, em seguida, rapidamente inserido no “*Agilent 2100 Bioanalyzer*” para início da eletroforese. Não é recomendado ultrapassar cinco minutos entre a montagem do

chip e o início da corrida para que amostras não sofram evaporação e a análise seja prejudicada.

3.8.2.4. Construção da biblioteca de DNA fita-simples

As amostras de DNA genômico foram utilizadas na construção de bibliotecas de fragmentos de DNA fita-simples (ssDNA) para o pirosequenciamento com o protocolo “*GS FLX Titanium Series - Rapid Library Preparation Method*” da 454 Roche. Amostras de 500ng de DNA genômico de *X. fastidiosa* (cepas Hib4 com 4 passagens em cultivo e Pr8x com 7 passagens em cultivo) purificado e previamente avaliado e quantificado como descrito acima foram transferidas para microtubos de 1,5 mL e o volume final para 100µL foi completado com tampão TE. As amostras foram então depositadas no fundo do aparato de nebulização (*nebulizer cup*) seguindo-se adição de 500µL do tampão de nebulização e homogeneização da mistura.

A nebulização foi realizada em fluxo laminar com pressão de nitrogênio gasoso de 30 psi (2,1 bar) e com tempos que variaram de acordo com a amostra. Para a amostra de Pr8x o tempo de nebulização foi de 1 minuto e 10 segundos e para a amostra de Hib4, foi de 2 minutos. A fragmentação ideal gera fragmentos de 500 a 800 pares de base.

Após a nebulização, 2,5 mL de tampão PBI (*MinElute PCR Purification Kit*, Qiagen) foram adicionados à amostra no “*nebulizer cup*”, seguindo-se homogeneização e posterior filtração em colunas *MinElute*. A purificação consistiu na adição sucessiva de 750 µL da amostra de DNA nebulizado em uma coluna de purificação e, a cada adição o tubo era centrifugado por 15 segundos a 13000 rpm. Após a passagem de toda a amostra, o tubo com a coluna foi centrifugado por um minuto a 13000 rpm. Em seguida, a coluna foi lavada por centrifugação a 13000 rpm por um minuto com 750 µL de tampão PE. Após nova centrifugação por mais um minuto, a coluna foi invertida e centrifugada por mais um minuto, a 13000 rpm. Os fragmentos purificados foram eluídos com 16 µL de tampão TE através de centrifugação por um minuto a 13000 rpm.

Em seguida, foi realizada a etapa de “*Fragment End Repair*” ou polimento das extremidades dos fragmentos pela adição de 2,5 µL de tampão PNK 10x, 2,5 µL de ATP, 1 µL de dNTP, 1 µL de T4 DNA polimerase, 1 µL de PNK (polinucleotídeo quinase) e 1 µL de Taq polimerase e incubação em termociclador (Applied Biosystems) por 20 minutos a 25°C e 20 minutos a 72°C. A reação foi armazenada a 4°C.

A etapa seguinte consiste na ligação dos adaptadores A e B às duas extremidades dos fragmentos com DNA ligase. A reação consistiu na incubação dos reagentes (ligase e adaptadores) com a amostra de fragmentos purificada por 10 minutos a 25°C. Estes adaptadores são um par de oligonucleotídeos dupla-fita que possibilitam a imobilização dos fragmentos de DNA nas microesferas, além de gerar regiões iniciadoras para a PCR em emulsão (emPCR) e para o sequenciamento. Eles possuem ainda uma sequência palindrômica de quatro nucleotídeos que será utilizado pelo programa de análise de dados do sequenciador 454 para reconhecer e selecionar os *reads*. O adaptador B possui biotina na sua extremidade 5' que garante a imobilização da biblioteca de dsDNA nas *beads* de imobilização (que possuem estreptavidina acoplada) e o subsequente isolamento da biblioteca de ssDNA que servirá de molde para o sequenciamento. Os adaptadores A e B não fosforilados são ligados ao DNA em reações que contém excesso molar de adaptadores para maximizar sua ligação ao DNA dupla-fita e evitando a formação de concatâmeros.

Em seguida, os fragmentos foram recuperados pelo método *Double SPRI* (*Solid Phase Reversible Immobilization*) que consiste na utilização de *beads* magnéticas (*AMPure beads*, Agencourt AMPure System, Beckman Coulter) que se ligam preferencialmente a fragmentos de DNA de tamanhos maiores. A faixa de tamanho desejada dos fragmentos é de 400 a 800pb, com tamanho médio de ~700pb. Após todo o processo de ligação dos fragmentos às *beads*, lavagens são realizadas para a retirada dos fragmentos que não contém os adaptadores e por isso não ligaram às *beads*. As *beads* com os fragmentos são lavadas com etanol 70%, incubadas a 37°C por um minuto. A biblioteca de ssDNA é ressuspensa em 50µL de tampão TE e separada das *beads* magnéticas utilizando o *Magnetic Particle Concentrator*. A biblioteca de ssDNA é recuperada no sobrenadante e será então submetida a emPCR com *primers* complementares aos adaptadores A e B.

Após obtenção da biblioteca purificada é necessário quantificar o número de moléculas de DNA e analisar a distribuição dos tamanhos dos fragmentos de DNA. A quantificação é realizada em placas de 96 poços e lida em espectrofluorímetro. Uma curva padrão é confeccionada a partir da leitura de uma diluição seriada de uma amostra de DNA fornecida pelo fabricante. Em cada poço, deve-se adicionar 50µL das amostras, sendo que no poço com a amostra da biblioteca a ser medida deve-se aplicar todo o seu conteúdo. Após a medição, a biblioteca é recuperada e transferida para um novo tubo. A curva padrão e o cálculo da concentração da biblioteca foram obtidos com o uso da "*Rapid Library Quantitation Calculator*"

disponível em www.454.com/my454 e amostra é diluída para 1×10^7 moléculas/ μL . A análise da distribuição de tamanhos dos fragmentos na biblioteca foi realizada no equipamento *Agilent 2100 Bioanalyzer*, utilizando o *High sensitivity DNA chip* (descrito no item 3.8.2.3) com $1 \mu\text{L}$ da biblioteca.

3.8.2.5. Titulação da biblioteca e PCR em emulsão em pequena escala

Uma vez quantificada e analisada quanto a sua qualidade, a biblioteca de DNA é utilizada para a amplificação dos fragmentos por PCR em emulsão. A reação de emPCR requer que cada *bead* esteja ligada a apenas uma molécula molde de ssDNA amplificável. Para garantir o sucesso do sequenciamento, a quantidade ótima de DNA adicionada à solução de *beads* deve ser determinada uma vez que, se pouco DNA é adicionado, a reação de emPCR resultaria em um grande número de *beads* que não estariam ligadas a DNA molde. Por outro lado, se for adicionado DNA em excesso, algumas *beads* conteriam múltiplos moldes amplificáveis o que é inadequado para o sequenciamento. A quantidade de biblioteca de ssDNA que resulta em uma única e efetiva molécula de DNA molde ligada a cada *bead* depende da qualidade do DNA, da composição da biblioteca, bem como de outros fatores não identificados.

Inicialmente foi realizada a titulação da biblioteca utilizando o protocolo de emPCR em pequena escala, “*GS FLX Titanium Series emPCR Method– Lib-L SV*” e o “*GS FLX Titanium Small Volume emPCR Kit*” para determinar a porcentagem de *beads* enriquecidas (o valor ideal é de 8% de *beads* enriquecidas). Para tal, volumes diferentes de cada biblioteca de ssDNA na concentração de 1×10^6 e 2×10^6 moléculas de DNA/ μL (diluídas da amostra com 1×10^7 moléculas de DNA/ μL) foram misturados a um número fixo de *capture beads* de modo a obtermos 0,5, 2, 4, 8 cópias por *bead*.

Nesta etapa, os fragmentos que constituem a biblioteca de ssDNA são imobilizados nas *capture beads* para permitir sua separação na emulsão. O objetivo é garantir que cada *capture bead* esteja ligada a apenas uma molécula de DNA e que cada microreator aquoso possua apenas uma *capture bead*. A imobilização às *capture beads* envolve o anelamento do adaptador B, que está ligado à extremidade de cada molécula molde que constitui a biblioteca, aos iniciadores de captura imobilizados nessas *beads*. Para reação de emPCR, as *capture beads* com a molécula de ssDNA imobilizada são ressuspensas com os reagentes de amplificação e óleo, seguindo-se vigorosa agitação mecânica no equipamento *TissueLyser* (Qiagen) a 25 Hz por 2 minutos, para formação da emulsão.

Após a emulsificação, a mistura para amplificação contendo as *beads* com fragmentos foi distribuída em uma placa de 96 poços, sendo 100µL em cada poço. A emPCR seguiu-se com o seguinte programa de amplificação: 94°C por 4 minutos, 50 ciclos de 94°C por 30 segundos, 58°C por 4,5 minutos e 68°C por 30 segundos, sendo mantidas ao final a 10°C até serem retiradas do termociclador. As emulsões e os microreatores foram então quebrados, para a recuperação das *beads* contendo a biblioteca de ssDNA amplificada, com o auxílio de uma seringa, seguindo-se diversas lavagens com isopropanol.

O processo de amplificação gera uma proporção de *beads* cuja amplificação não foi bem sucedida. Desta forma, o passo seguinte consiste no enriquecimento (separação) da população de *beads* da biblioteca de ssDNA efetivamente amplificada. Esta etapa baseia-se no anelamento do *primer* de enriquecimento, o qual é biotinizado, ao adaptador A de cada molde amplificado, e a posterior ligação desse *primer* a *beads* magnéticas contendo estreptavidina. As *beads* contendo a biblioteca de ssDNA amplificada são então separadas dessas *beads* magnéticas por desnaturação dos produtos de amplificação do *primer* de enriquecimento utilizando NaOH. A proporção final entre a quantidade de *beads* obtidas com o enriquecimento e a quantidade inicial de *beads* utilizada na emPCR corresponde à porcentagem de *beads* enriquecidas e é um indicativo da qualidade da biblioteca amplificada. Para a contagem de *beads*, 3 µL das amostras são adicionadas a 10 mL de Isoton e avaliadas no equipamento *Beckman Coulter*. Com os valores das porcentagens de enriquecimento em cada diluição da titulação, foram construídas as curvas de titulação para ambas as amostras e a quantidade ótima de biblioteca de DNA a ser utilizada na emPCR em grande escala é aquela que resulta na porcentagem de 8% de *beads* enriquecidas, segundo o fabricante. Este enriquecimento corresponde à quantidade que resultou, supostamente, em uma única molécula de DNA utilizada como molde para a amplificação.

3.8.2.6. emPCR em grande escala ou definitiva

Para emPCR em grande escala de Pr8x e Hib4, escolhemos a quantidade da biblioteca de ssDNA equivalente a, respectivamente, 4 cópias e 5 cópias por *bead*. De acordo com os gráficos de titulação obtidos como descrito acima (não mostrados), estes valores corresponderiam a aproximadamente 10-11% de enriquecimento ao final do processo. Este valor tem se mostrado mais adequado para compensar perdas que ocorrem durante o processo de emPCR em larga escala (Dra. Layla Farage Martins, comunicação pessoal).

O procedimento para realização da emPCR em larga escala se assemelha ao procedimento de emPCR em pequena escala, exceto pelo maior volume inicial. O processo de ligação às *capture beads*, formação da emulsificação, mistura de amplificação e reação de amplificação são os mesmos que o processo de emPCR em pequena escala. A reação é realizada em *cups* e não mais microtubos. Ao final do processo, cada biblioteca gerará duas placas de 96 poços inteiras para a etapa de amplificação.

Após a amplificação, a emulsão e os microreatores são quebrados para separação das *beads* amplificadas, neste caso com um aparato interligado a dois tubos de 50 mL ligado a uma bomba à vácuo, e não manualmente com uma seringa. Novamente as lavagens são realizadas diversas vezes com isopropanol. O processo de enriquecimento das *beads* assemelha-se a emPCR em pequena escala descrito acima, incluindo a contagem de *beads* no equipamento *Beckman Coulter* para estimativa do enriquecimento obtido. Espera-se que a porcentagem de enriquecimento após a emPCR em larga escala esteja entre 5% e 20%.

Entretanto, ambas as bibliotecas tiveram suas porcentagens de enriquecimento após emPCR definitiva acima dos valores ideais: a porcentagem de enriquecimento para a cepa Pr8x foi de 45% e para a cepa Hib4 de 28%. Estes altos valores de enriquecimento significam uma alta probabilidade de existirem *beads* mistas, ou seja, que contenham mais de um fragmento de ssDNA ligados a uma única *bead* antes do processo de emPCR. Este fato provavelmente pode ter ocorrido devido a uma maior porcentagem de cópias por *bead* que utilizamos para iniciar a emPCR em larga escala e de que perdas previstas não ocorreram. Ainda assim, optamos por prosseguir com o pirosequenciamento. Durante o processo de sequenciamento, os poços contendo as *beads* mistas gerarão dois sinais distintos simultaneamente, uma vez que existem mais de um fragmento de DNA ligado à *bead*; desta forma, no processamento das imagens, o próprio *software* do equipamento 454/Roche excluirá tais poços. Portanto, um número menor de sequências será obtido. Por outro lado, com o alto enriquecimento, a quantidade de *beads* é suficiente para realizar duas corridas de pirosequenciamento.

3.8.2.7. Preparação da placa e pirosequenciamento

A etapa de sequenciamento é dividida em três fases: o preparo da *PicoTiterPlate* (PTP), a pré-lavagem do sequenciador e a corrida de sequenciamento. É importante sincronizar o preparo da placa e do sequenciador, para assegurar que a corrida de sequenciamento seja iniciada imediatamente após o término do

preparo da placa. O aparelho *Genome Sequencer FLX* deve ser cuidadosamente lavado antes de cada corrida de sequenciamento.

Após a amplificação, as *beads* recobertas com fragmentos de DNA são depositadas e imobilizadas na PTP que contem 1,6 milhões de poços, os quais têm a dimensão exata para comportar uma única microesfera. A PTP foi previamente carregada com outras *beads* que garantem o empacotamento e a formação de camadas de apoio. Nas duas corridas que realizamos, 2 milhões de *beads* das bibliotecas de ssDNA das cepas Hib4 e Pr8x foram carregadas separadamente em duas regiões da PTP. Em seguida são aplicadas as *beads* ligadas a enzimas que atuam no processo de pirosequenciamento e a placa pronta é inserida no sequenciador, seguindo-se processamento automatizado para adição sucessiva dos reagentes de sequenciamento.

A reação de pirosequenciamento ocorre em ciclos, paralelamente em cada um dos milhões de poços da PTP, sendo que cada um deles contem, idealmente, uma única *bead* com moléculas de ssDNA idênticas resultantes da etapa de amplificação na emPCR. A reação é iniciada a partir de um *primer* de sequenciamento que é anelado a extremidade de cada molécula de DNA imobilizada nas *beads*. A cada ciclo, a placa recebe sequencialmente soluções de dNTPs (dGTP, dTTP, dCTP e dATP), a DNA polimerase adiciona o nucleotídeo complementar ao molde, pirofosfato é liberado e utilizado pela ATP sulfurilase para síntese de ATP, o qual é hidrolisado pela luciferase na reação com luciferina para produção de luz. Segue-se a eliminação do dNTP residual pela enzima apirase e, liberação do fluxo com o dNTP seguinte. Desta forma a cada ciclo, uma câmera CCD captura imagens para cada fluxo de um dos 4 dNTPs, onde pontos luminosos significam que ocorreu a incorporação do nucleotídeo naquele poço da placa.

Os dados brutos de uma corrida de sequenciamento consistem de uma série de imagens digitais capturadas pela câmera CCD, representando a superfície da PTP durante as reações de sequenciamento em tempo real. A quantidade de luz emitida é proporcional ao número de nucleotídeos que são adicionados pela DNA polimerase. Em seguida, os dados brutos serão filtrados quanto à presença de *beads* mistas e os dados brutos serão processados. *Beads* mistas são detectadas pelo próprio sequenciador a partir da presença de dois ou mais sinais luminosos distintos e simultâneos no mesmo poço durante o sequenciamento o que significa a presença de mais de um fragmento de ssDNA ligado a uma *bead*, resultado de uma diluição errada da biblioteca para a realização da emPCR em larga escala.

3.8.2.8. Obtenção e processamento dos dados brutos

Após ~10-12 horas de corrida, são obtidos dados brutos das imagens dos sinais luminosos gerados a partir da *PTP* e detectados por uma câmera CCD. Após ~10 horas adicionais de análise computacional, as imagens são processadas pelo *software* que acompanha o 454 *Genome Sequencer*, e convertidas em sequências (*reads*), cada uma derivada das imagens de um poço da placa. O processamento dos dados consiste de uma série de etapas de correção automática dos dados para compensar efeitos óticos e alguma ineficiência química do processo. As sequências processadas são então submetidas a filtros de qualidade que identificam e excluem aquelas que apresentam baixa qualidade. As sequências de alta qualidade são submetidas a um algoritmo que converte o sinal de fluorescência em sequência de bases e atribui a cada base um *score* de qualidade ou confiabilidade. Com a versão dos kits e reagentes que utilizamos, são obtidas aproximadamente 1 milhão de sequências, cada uma com mais de 400 bases com 99% de confiança (*Phred score* > 20). Os dados brutos, metadados e dados métricos são armazenados em planilhas e arquivos texto, a partir dos quais informações sobre o sequenciamento podem ser extraídas. Um dos tipos de arquivos que é gerado contém a sequência de cada *read* associada à qualidade de cada base, o qual é utilizado como entrada nos programas de montagem dos *contigs*, como o *Newbler*.

O processo do pirossequenciamento foi realizado no CATG (Centro Avançado de Tecnologia em Genômica) do Instituto de Química/USP (IQUSP) segundo as recomendações dos protocolos da 454 Life Sciences/Roche, com o apoio da Dra. Layla Farage Martins (técnica responsável pelo CATG).

3.9. Análises bioinformáticas

3.9.1. Montagem e anotação de genomas

A montagem de *contigs* e *scaffolds* foi realizada *ab initio* com as sequências obtidas no sequenciador 454 utilizando-se os Programas *Newbler* (454 Life Sciences/Roche), MIRA (http://chevreux.org/projects_mira.html) e Cross_Match (<http://www.phrap.org/phredphrap/general.html>) aliados a cuidadosas intervenções manuais. As montagens foram realizadas pelo Dr. João Carlos Kitajima. Algumas etapas de verificação das montagens contaram com o apoio de Luiz Gonzaga Paula de Almeida (Laboratório Nacional de Computação Científica, Petrópolis, RJ).

A anotação automática e análises comparativas das CDS/ORFs dos genomas de *X. fastidiosa* foram realizadas utilizando-se o servidor RAST (*Rapid Annotation using Subsystem Technology*) (Aziz *et al.*, 2008) e IMG/ER (Integrated Microbial Genomes-Expert Review (Markowitz *et al.*, 2012), respectivamente, disponíveis em <http://rast.nmpdr.org/> e <http://img.jgi.doe.gov/>. Estas plataformas fazem a predição de genes e sua anotação automática integrando informações e termos de bancos de dados tais como Nomenclatura de Enzimas (Bairoch, 2000), COG (Tatusov *et al.*, 2003), Pfam (Punta *et al.*, 2012), TIGRfams (Selengut *et al.*, 2007), InterPro (Hunter *et al.*, 2012) e *Kegg Ortholog* (Kanehisa *et al.*, 2012), entre outros. Além disso, tanto RAST como IMG/ER disponibilizam diversas funcionalidades para análises comparativas diversas como BLAST (*Basic Local Alignment Search Tool*) (Altschul *et al.*, 1990; Altschul *et al.*, 1997), Artemis (Carver *et al.*, 2012) e MUMmer (Delcher *et al.*, 2003), além da visualização gráfica de mapas cromossômicos (*Genome browser*) e ferramentas para anotação manual pelo usuário.

As análises comparativas da anotação também foram realizadas em bases públicas de dados tais como *GenBank* do NCBI (*National Center for Biotechnology Information*, <http://www.ncbi.nlm.nih.gov/>), *Xylella fastidiosa* *Genome Project* (<http://aeg.lbi.ic.unicamp.br/xf/>) e *Xylella fastidiosa*/Pierce's Disease Strain *Genome Project* (<http://bioinfo03.ibi.unicamp.br/lbi/index.html>).

3.9.2. Análises filogenéticas

Para as construções de árvores filogenéticas foi utilizado o método probabilístico de Máxima Verossimilhança (ML), com o modelo geral de tempo-reverso (GTR + G).

A reconstrução filogenética realizada neste trabalho foi feita com o método de Análise de Sequência Multilocus (MLSA). A MLSA é um método de reconstrução filogenética que possibilita a obtenção de uma classificação taxonômica mais robusta de cepas proximamente relacionadas e com alto nível de recombinação genética (Maiden *et al.*, 1998; Hanage *et al.*, 2005). O princípio básico deste método consiste essencialmente na concatenação de várias sequências para o mesmo conjunto de organismos, resultando em uma "supermatriz", que é utilizada para inferir a filogenia (Maiden *et al.*, 1998). Foram utilizados sequências de 7 genes *housekeeping*, utilizados em um trabalho recente com sequências de cepas de *X. fastidiosa* (Nunney *et al.*, 2012), os quais são: *leuA* (XF1818), *petC* (XF0910), *malF* (XF2447), *cysG* (XF0832), *holC* (XF0136), *nuoL* (XF0316) e *glfT* (XF0656). A identificação entre parênteses refere-se a CDS no genoma da cepa 9a5c que

codificam, respectivamente, 2-isopropilmalato sintase, ubiquinol citocromo C oxidorreductase, permease transportadora de açúcar do tipo ABC, siroheme sintase, subunidade chi da DNA polimerase III, NADH-ubiquinona oxidorreductase e proteína de simporte de glutamato. As sequências dos genes concatenados somou aproximadamente 8.500 pb. As sequências completas destes genes para todas as cepas incluídas na filogenia foram recuperadas de bases de dados públicas, como o GenBank, e de dados gerados pelo sequenciamento de genomas realizado neste trabalho e em trabalhos que estão sendo realizados em nosso grupo (Santana *et al.*, 2010). As sequências concatenadas foram alinhadas utilizando-se o *software* Bioedit (Hall, 2005). A partir dos alinhamentos foram calculadas as árvores filogenéticas utilizando-se o *software* PAUP4b10 (Wilgenbusch e Swofford, 2003). As sequências dos genes de *Xanthomonas albilineans* foram usadas para gerar o grupo externo (Pieretti *et al.*, 2009).

3.9.3. Outras análises bioinformáticas

Para o alinhamento de sequências de aminoácidos ou de nucleotídeos utilizamos os programas ClustalW2 (Thompson *et al.*, 2002) ou BioEdit (Biological Sequence Alignment Editor) (Hall, 2005).

O planejamento de oligonucleotídeos foi realizado com o programa PRIMER3 (Rozen e Skaletsky, 2000).

3.10. Ensaios de infecção em cafeeiros com *X. fastidiosa*

3.10.1. Inoculação da suspensão bacteriana

Plantas de café com aproximadamente 2 meses de crescimento, foram inoculadas com suspensões de células de *X. fastidiosa* cepa J1a12 (isolada de *Citrus sinensis* com sintomas de CVC) (Monteiro *et al.*, 2001) como cepa teste e, com suspensões de células da cepa 3124 (isolada de *Coffea arabica* com sintomas de requeima de café) (Li *et al.*, 2001) como controle positivo.

Para a preparação do inóculo, uma alçada de células foi removida de uma placa de cultivo em meio PWG-ágar, crescidas por sete dias, e transferidas para um tubo contendo 500 µL de tampão PBS. Após homogeneização da suspensão de células, foram preparadas diluições seriadas (10^{-1} , 10^{-2} , 10^{-3} e 10^{-4}), sendo as três últimas diluições semeadas, por gotejamento, em placas de meio BCYE (Wells *et al.*, 1981), para cálculo das unidades formadoras de colônia (UFC). As densidades óticas a 600nm (DO_{600}) dos homogeneizados de células foram também aferidas.

As suspensões de células (não diluídas) foram diretamente utilizadas na inoculação de plantas de café, sendo 25 plantas inoculadas com as células da cepa J1a12 com 24 passagens e 25 plantas com as células da cepa 3124 com 15 passagens. As inoculações foram feitas com a adição de 10 μ L do homogeneizado de células na junção de três folhas, próxima a região apical do caule, seguida de perfurações cuidadosas e superficiais com uma agulha fina, para facilitar a penetração das bactérias no tecido vegetal (Figura 1). As plantas foram privadas de água por alguns dias anteriores à inoculação para captação da suspensão bacteriana. Os controles negativos foram uma planta inoculada apenas com tampão PBS e uma planta não-inoculada.

Figura 1. Detalhe das plantas de cafeeiros inoculadas e do processo de inoculação. As plantas inoculadas tinham aproximadamente 2 meses de crescimento e estavam há alguns dias submetidas à privação de água. O procedimento de inoculação foi realizado com a aplicação da suspensão de bactérias celular no caule da planta com subsequentes perfurações superficiais com auxílio de uma agulha.

Todos os vasos com as plantas de café foram devidamente identificados com placas de plástico ao redor do caule da planta e as plantas foram armazenadas em estufa climatizada onde são regadas periodicamente. Este experimento está sendo realizado no Centro de Citricultura Sylvio Moreira (Cordeirópolis, SP) sob supervisão do Dr. Helvécio Della-Colleta Filho. O aparecimento de sintomas é

visualmente inspecionado e a colonização da bactéria é aferida através de ensaios de PCR.

3.10.2 Extração de DNA do pecíolo de plantas inoculadas com *X. fastidiosa*

Na tentativa de avaliar a colonização das cepas J1a12 e 3124 de *X. fastidiosa* inoculadas em plantas de cafeeiro, foram realizados procedimentos de extração de DNA do tecido vegetal, mais precisamente dos pecíolos e da nervura central das folhas coletadas. Em uma primeira extração, três meses após inoculação, foram coletadas folhas de diferentes regiões das plantas (diferentes alturas), de modo a obter-se um perfil geral de colonização ao longo de toda planta. Das 25 plantas inoculadas com cada cepa, 12 plantas foram usadas neste primeiro experimento de extração. Além disso, foram usadas amostras de folhas das plantas usadas como controle negativo (inoculada somente com tampão PBS ou não-inoculada). Em uma segunda extração, oito meses após inoculação, outras 12 plantas tiveram suas folhas retiradas para a análise do perfil de colonização destas cepas por toda a planta. Novamente, folhas das plantas usadas como controle negativo também foram retiradas para extração de DNA.

Em ambos os processos de extração, as folhas foram retiradas e mantidas no gelo até a sua manipulação, de modo a evitar que ocorra a oxidação do material vegetal, uma vez que plantas de café contêm uma maior quantidade de compostos fenólicos (Helmécio Della-Coletta Filho, comunicação pessoal).

Para extração de DNA do tecido vegetal, foi utilizado um protocolo descrito por (Murray e Thompson, 1980). As amostras de tecido vegetal de cada folha foram obtidas a partir do pecíolo até a metade da nervura central. Este material foi cortado em pedaços menores e depositados em tubos de 2 mL contendo duas *beads* metálicas (*Stainless Steel Beads*, Qiagen). Em seguida foram adicionados 1,4 mL do tampão de extração CTAB contendo sarcosil e β -mercaptoetanol (Tris.HCl 1M pH7,5, 100 mL/L; EDTA 0,5 M, 20 mL/L; NaCl 5 M, 140 mL/L; CTAB 5%, 200 mL/L; Sarcosyl 10%, 200 mL/L; β -mercaptoetanol 140 mM, 10 mL/L; após mistura dos reagentes, o tampão deve ser incubado a 60°C para total homogeneização). Após a adição do tampão, o material vegetal foi triturado utilizando o equipamento *TissueLyser* (Qiagen), com três etapas de trituração de 30 segundos cada à uma frequência de 30 Hz, sendo os tubos mantidos em gelo nos intervalos.

Após trituração do material vegetal, o homogeneizado em cada tubo de amostra foi transferido em iguais volumes para dois tubos de 1,5 mL. Os tubos

foram incubados a 60°C em banho-maria durante 30 minutos. As amostras foram agitadas periodicamente durante este período. Após este intervalo, os tubos foram retirados do banho-maria e deixados esfriar a temperatura ambiente. Em seguida, foi adicionado um volume da solução de clorofórmio:álcool isoamílico (24:1); os tubos foram agitados levemente por volta de 80 vezes. As amostras foram centrifugadas a 10000 rpm durante 8 minutos e a fase superior da suspensão foi coletada e transferida para outro tubo de 1,5 mL. Em seguida, foi adicionado 0,1 volume de CTAB 10% (CTAB, 100 g/L; NaCl 5 M, 140 mL/L). Novamente, um volume de clorofórmio:álcool isoamílico (24:1) foi adicionado, seguindo agitação suave e centrifugação a 10000 rpm por 8 min. A fase superior de cada tubo foi transferida para outro tubo de 1,5 mL. O próximo passo corresponde à adição de um volume do tampão de precipitação CTAB (CTAB, 10 g/L; Tris.HCl 1 M pH7,5, 50 mL/L; EDTA 0,5 M, 20 mL/L), seguindo agitação suave e incubação por 15 min à temperatura ambiente. As amostras foram centrifugadas por 4 min a 10000 rpm e o sobrenadante foi descartado. Foram adicionados 0,4 mL de TE Alto Sal (Tris.HCl 1 M pH7,5, 10 mL/L; EDTA 0,5 M, 2 mL/L; NaCl 5 M, 200 mL/L) e os tubos foram incubados em banho-maria à 65°C até total dissolução do precipitado. Em seguida, foram adicionados 2 volumes de etanol absoluto gelado, seguindo-se centrifugação por 5 min a 12000 rpm. O sobrenadante foi descartado e o precipitado foi lavado com 0,7 mL de etanol 70% e centrifugado por 5 min a 12000 rpm. O sobrenadante foi descartado e o precipitado foi novamente lavado com 0,4 mL de etanol 100% e centrifugado por 5 min a 12000 rpm. O sobrenadante foi removido, o precipitado seco em câmara de fluxo laminar, ressuspendido em 25 µL de TE 1/10 com RNAase (Tris.HCl 1 M pH7,5, 1 mL/L; EDTA 0,5 M, 0,2 mL/L) e deixado por algumas horas (2-3h) em banho-maria a 65°C para completa solubilização do DNA. A quantificação da concentração e qualidade do DNA foi avaliada como descrito no item 3.3.

As amostras de DNA extraídas foram analisadas por PCR quanto à presença de DNA de *X. fastidiosa* com o par de *primers* CVC-1/272-2-int (Pooler e Hartung, 1995) como descritos no item 3.5.1.

3.10.3. Recuperação e isolamento de *X. fastidiosa* de plantas de café

Cinco meses após a inoculação das cepas J1a12 e 3124 em plantas de café, foi realizado o procedimento para recuperação e isolamento de *X. fastidiosa*. Para tal as mesmas plantas que foram utilizadas no ensaio de extração de DNA tiveram suas folhas removidas e toda a parte aérea do caule foi cortada, desde a base até o topo. Os caules foram lavados com detergente e água corrente e transferidos para

câmara de fluxo laminar, onde foram cortados em pedaços menores e externamente esterilizados pela lavagem sequencial por 2 min em hipoclorito de sódio, água destilada estéril e etanol 70%, seguindo duas lavagens adicionais em água destilada estéril. Após secagem no interior da câmara de fluxo laminar, os pedaços de caule secos foram espremidos com o auxílio de um alicate para extração das seivas dos vasos xilemáticos e floemáticos. O líquido assim obtido foi semeado em meio PWG-ágar por gotejamento e as placas foram então incubadas a 28°C por ~ 20 dias, porém nenhuma colônia de *X. fastidiosa* foi observada.

4. Resultados e Discussão

4.1. Avaliação fenotípica de *Xylella fastidiosa* cepas J1a12, Pr8x e Hib4

As cepas de *Xylella fastidiosa* J1a12, Pr8x e Hib4 selecionadas para o sequenciamento genômico foram isoladas no Estado de São Paulo a partir de hospedeiros sintomáticos. A tabela 3 apresenta as localidades e os hospedeiros destas três cepas e da cepa 9a5c, considerada a cepa de referência neste estudo.

Tabela 3. Descrição das cepas de *X. fastidiosa* selecionadas para seqüenciamento

Cepa	Hospedeiro	Origem	Referência
9a5c	<i>Citrus sinensis</i> L. Osbeck cv. Natal	Macaubal, SP, Brasil	(Li <i>et al.</i> , 1999)
J1a12	<i>Citrus sinensis</i> L. Osbeck cv. Pêra	Jales, SP, Brasil	(Monteiro <i>et al.</i> , 2001)
Pr8x	<i>Prunus salicina</i>	Jarinu, SP, Brasil	Presente estudo ^a
Hib4	<i>Hibiscus rosa-sinensis</i>	Campinas, SP, Brasil	Presente estudo ^a

^aAs cepas Pr8x e Hib4 foram cedidas pelo Dr. Helvécio Della Colleta Filho, Centro de Citricultura Sylvio Moreira, Cordeirópolis, SP

O cultivo destas cepas está estabelecido em laboratório, podendo ser realizado em meio PWG (Davis, 1981) ou meio BCYE (Wells *et al.*, 1981). A figura 2 ilustra a cultura das cepas 9a5c, J1a12, Hib4 e Pr8x após 7 dias de crescimento em meio PWG. Como esperado, a cepa J1a12 forma agregados menos compactos do que a cepa 9a5c, apresentando maior proporção de células planctônicas (Figura 2B). Análises anteriormente realizadas, utilizando-se hibridização em microarranjos de DNA (Koide *et al.*, 2004) e sequenciamento preliminar de seu genoma (da Silva, 2010), sugerem que o fenótipo menos agregativo da cepa J1a12 é provavelmente resultado da ausência de uma cópia do gene da adesina da fimbria (pilus do tipo I). Este fenótipo menos agregativo parece estar associado a sua ineficácia em colonizar o hospedeiro vegetal e seu fenótipo avirulento em plantas de citros e tabaco (Koide *et al.*, 2004; da Silva, 2010).

A cepa Pr8x apresenta um fenótipo muito semelhante à cepa 9a5c no cultivo em meio líquido, tanto em relação formação de agregados mais compactos, como na formação de biofilme aderido à superfície do frasco de cultivo (Figura 2C).

A cepa Hib4 apresenta um fenótipo de agregação diferente das demais cepas, formando agregados com tamanho intermediário entre aqueles observados em 9a5c (agregados celulares maiores e maior quantidade de biofilme) e J1a12 (agregados

celulares menores e maior quantidade de células planctônicas). Além disso, Hib4 mostra menor aderência ao vidro se comparada às outras cepas, sendo que praticamente não se observa o anel do biofilme (Figura 2D).

Figura 2. Cultivo de cepas de *X. fastidiosa*. A figura ilustra o cultivo das cepas 9a5c (A), J1a12 (B), Pr8x (C) e Hib4 (D) em meio PWG líquido após 7 dias de crescimento a 28°C com agitação de 100 rpm.

Os fenótipos das cepas 9a5c, J1a12, Pr8x e Hib4 também foram analisados no cultivo em meio PWG sólido. As características das colônias destas cepas foram observadas na microscopia ótica e estão resumidas na Tabela 4.

Tabela 4. Características morfológicas das colônias das cepas de *X. fastidiosa*

Morfologia das colônias	Cepas	9a5c	J1a12	Pr8x	Hib4
	Colônias	Definidas	Dispersas	Definidas	Pouco definidas
	Forma	Circular	Irregular	Circular	Irregular
	Margem	Ondulada	Indefinida	Ondulada	Indefinida
	Aparência da superfície	Rugosa	Lisa	Rugosa	Rugosa
	Textura	Áspera/seca	Viscosa	Áspera/seca	Viscosa
	Elevação	Convexa	Achatada	Convexa	Convexa
	“Franja”	Sim	Sim	Não	Sim

Observamos que as cepas 9a5c e Pr8x apresentam características morfológicas similares quanto à forma e textura de suas colônias. Formam grandes agregados de células, além de apresentarem uma textura áspera e seca. Porém, enquanto as colônias da cepa 9a5c exibem uma pequena “franja” em seu entorno e presença de colônias satélites indicativos da ocorrência de *twitching motility*, as colônias de Pr8x não exibem esta “franja” nas condições analisadas (dados não mostrados). Assim como em outras bactérias, em *X. fastidiosa*, a motilidade *twitching* depende do pilus do tipo IV (Mattick, 2002; Meng *et al.*, 2005). Portanto, a cepa Pr8x ou apresenta um defeito na montagem desta estrutura ou não expressa este tipo de pilus ou movimento nas condições avaliadas.

Também em meio sólido, a cepa J1a12 apresentou um fenótipo de colônias mais dispersas formando agregados muito menores comparados com as colônias de 9a5c e Pr8x. Além disso, suas colônias são viscosas em comparação com as colônias com aspecto áspero de 9a5c e Pr8x. Por outro lado, assim como no caso de 9a5c, foi possível observar a “franja” e colônias satélites afastadas da colônia principal na cultura de J1a12 (dados não mostrados).

A cepa Hib4 apresenta um fenótipo intermediário às cepas 9a5c e J1a12, assim como observado no cultivo em meio líquido. Suas colônias formam agregados menores que os agregados observados na cepa 9a5c, porém maiores que os agregados formados na cepa J1a12. Suas colônias apresentam uma forma irregular além de certa viscosidade, aproximando-se mais ao fenótipo observado em colônias de J1a12.

4.2. Genotipagem de diferentes cepas de *X. fastidiosa*

Além de exibirem fenótipos diferenciados no cultivo em meio líquido ou sólido, as cepas 9a5c, J1a12, Pr8x e Hib4 também apresentam diferenças em ensaios de PCR com os pares de *primers* XF0077 R/F e XF1250 R/F, como mostrado na Figura 3A. Esses *primers* amplificam as CDS da adesina de fimbrias (XF0077) e da arginina deaminase (XF1250), gerando, respectivamente, produtos de 1599 pb e de 1122 pb para a cepa 9a5c. A cepa J1a12 exhibe um padrão de bandas distinto (549 pb para os *primers* referentes a CDS XF0077 e 159 pb para os *primers* da CDS XF1250) em razão de deleções previamente identificadas nesta região de seu genoma (Koide *et al.*, 2004). A PCR da cepa Pr8x gerou produtos de ~3000 pb (*primers* XF0077 R/F) e de 159 pb (*primers* XF1250 R/F) e portanto distinto dos padrões observados para 9a5c e J1a12. Entretanto, a cepa Hib4 apresentou um

padrão de bandas idêntico ao observado para a cepa J1a12 (549 pb para a CDS XF0077 e 159 pb para a CDS XF1250).

Um novo par de *primers* (XF1952 R/F) para genotipagem da cepa Hib4 foi definido, a partir das análises da sequência genômica da cepa Hib4 que será apresentada adiante. Este par de *primers* resulta na amplificação de uma região interna da CDS XF1952 (codifica uma proteína quinase relacionada à quimiotaxia) que na cepa Hib4 apresenta uma inserção de 21 nucleotídeos. Como mostra a figura 3B, esta inserção foi confirmada somente na cepa Hib4, gerando o produto esperado de 947 pb. Demais cepas, incluindo cepas U24d e Fb7 (citros) além de 3124 (cafeeiro), não apresentam produto de PCR com o par de *primers* XF1952 R/F.

Estes três pares de *primers* têm sido utilizados rotineiramente em nosso laboratório para genotipagem e diferenciação das cepas 9a5c, J1a12, Pr8x e Hib4.

Figura 3: Genotipagem das cepas J1a12, 9a5c, Hib4 e Pr8x através de PCR.

DNA total isolado destas cepas foi submetido a PCR com os pares de *primers* XF0077 R/F e XF1250 R/F (A) ou XF1952 R/F (B). **(A)** Canaleta 1: Marcador tamanho 1kb (Fermentas); canaleta 2: controle negativo do ensaio de PCR (sem DNA molde); canaleta 3: cepa J1a12; canaleta 4: cepa 9a5c; canaleta 5: cepa Hib4; canaleta 6: cepa Pr8x. **(B)** Canaleta 1: Marcador tamanho 1kb (Fermentas); canaleta 2: controle negativo do ensaio de PCR (sem DNA molde); canaleta 3: cepa J1a12; canaleta 4: cepa 9a5c; canaleta 5: cepa Hib4; canaleta 6: cepa Pr8x; canaleta 7: cepa U24d; canaleta 8: cepa 3124; canaleta 9: cepa Fb7.

Verificamos também que o par de *primers* CVC-1/272-2-int (Pooler e Hartung, 1995) tradicionalmente utilizado para diferenciar cepas de citros, resulta na amplificação de um produto de 500pb também para as cepas Hib4 e Pr8x (Figura 4). A PCR da cepa 3124 (de cafeeiro) com este par de *primers* também resulta em um produto de 500pb (Santana *et al.*, 2010). Portanto, podemos afirmar que este par de *primers*, aparentemente, distingue cepas sul-americanas e não apenas cepas isoladas de citros.

Figura 4: Análise com *primers* CVC1/272-2-int das cepas J1a12, 9a5c, Hib4 e Pr8x. Os produtos resultantes de ensaios de PCR com os pares de *primers* CVC1/272-2-int a partir de DNA total extraído das cepas J1a12 (canaleta 3), Pr8x (canaleta 4), Hib4 (canaleta 5) e 9a5c (canaleta 6) foram submetidos a eletroforese em gel de agarose 1%. O marcador de tamanho 1kbp (Fermentas) foi aplicado na canaleta 1. O controle negativo da PCR (sem adição de DNA molde) foi aplicado na canaleta 2.

4.3. Sequenciamento do genoma da cepa J1a12

O sequenciamento do genoma da cepa J1a12 (cepa não-virulenta) foi iniciado em projeto anterior em nosso laboratório (da Silva, 2010) no qual foi realizado o pirossequenciamento de bibliotecas de *shotgun* (bibliotecas de fragmentos de 500-700 bp obtidos por nebulização do DNA total). Contudo, a qualidade destas corridas foi abaixo do desejável e as sequências obtidas não possibilitaram a montagem do genoma completo, ainda que tenha sido possível confirmar várias diferenças genômicas previamente identificadas entre os genomas das cepas J1a12 e 9a5c

(Koide *et al.*, 2004; da Silva, 2010) bem como identificar um novo plasmídeo na cepa J1a12 em adição ao já conhecido pXF51 da cepa 9a5c. Com o objetivo de concluir o sequenciamento e a montagem do genoma e finalizar as análises comparativas do genoma da cepa J1a12, foi realizada uma nova corrida de sequenciamento na plataforma 454/Roche, mas utilizando a metodologia de *Paired End Tags* (PET) (Fullwood *et al.*, 2009). Para tal, o DNA total purificado da cepa J1a12 foi submetido à fragmentação controlada e os segmentos de ~3kb foram utilizados na construção da biblioteca de PET, seguindo-se nova fragmentação e pirosequenciamento na Unidade de Genômica Computacional Darcy Fontoura no Laboratório Nacional de Computação Científica (LNCC)¹. O sequenciamento com a estratégia PET oferece vantagens para a montagem de genomas, uma vez que pares de sequências separados um número de bases pré-determinado (no nosso caso 3kb) são identificados por *tags* orientando a posterior montagem *in silico* do genoma.

É importante destacar que a qualidade do DNA total purificado é fundamental para o sucesso do pirosequenciamento, tanto na estratégia *shotgun* como na estratégia PET. Idealmente, as preparações de DNA para pirosequenciamento devem apresentar razões de A260nm/230nm e A260nm/280nm entre 1,8 e 2,0 além de integridade adequada. A figura 5 mostra a integridade do DNA total da cepa J1a12 (purificado utilizando o procedimento descrito no item 3.3) avaliada através de eletroforese em gel de agarose. Uma amostra de DNA sem muitas quebras é ideal para o processo de pirosequenciamento, uma vez que esta ainda será nebulizada para a preparação da biblioteca. A pressão de gás nitrogênio usada e o tempo de nebulização busca de gerar fragmentos de tamanho entre 600 e 700 pb, no caso da estratégia *shotgun*, ou 3 a 8 kpb, no caso da estratégia PET, e portanto uma amostra que já esteja muito fragmentada, ao ser nebulizada, sofrerá novas quebras gerando fragmentos muito pequenos que diminuiriam o rendimento de todo o processo.

O resultado do pirosequenciamento com a estratégia PET foi bastante satisfatório. O total de bases sequenciadas foi de 84,2 Mpb e descontando-se as sequências dos adaptadores obtivemos 74,7 Mpb. Assim, a cobertura estimada do genoma de J1a12 foi de ~28 vezes, considerando o genoma da cepa 9a5c de ~ 2,7 Mpb como referência.

¹O Centro Avançado de Tecnologia em Genômica (CATG) do IQUSP ainda não possui o equipamento (*Hydroshear*) para a fragmentação controlada do DNA necessária para preparação da biblioteca de *Paired End Tags*. Desta forma, o pirosequenciamento utilizando esta estratégia teve de ser realizado no LNCC (Petrópolis, RJ) que possui tal equipamento.

Figura 5: Análise eletroforética da amostra de DNA total da cepa J1a12 utilizada para pirosequenciamento. Canaleta 1: Marcador de tamanho de 1kb (Fermentas); canaleta 2: DNA total da cepa J1a12.

4.3.1. Características do genoma da cepa J1a12

A montagem das sequências (300.869 reads/tamanho médio de 380bp) obtidas a partir do pirosequenciamento *shotgun* realizado anteriormente resultou em 374 *contigs* (maior *contig* com 115.071 pb) somando 2.614.884 pb (da Silva, 2010). Estas sequências foram reunidas com as sequências obtidas no pirosequenciamento com a estratégia *paired-end* e submetidas à montagem utilizando-se principalmente o programa Newbler que acompanha a plataforma 454/Roche. Esta montagem resultou em 233 *contigs* (maior *contig* com 212.828 pb), organizados em 18 *scaffolds* (maior *scaffold* com 523.148 pb) somando 2,78 Mpb e cobertura de ~65x. Apesar da elevada cobertura, a montagem do genoma de *X. fastidiosa* é complicada pela existência de regiões repetidas do genoma, principalmente as regiões dos pró-fagos (Van Sluys *et al.*, 2003; de Mello Varani *et al.*, 2008). Além disso, o tamanho médio das sequências obtido esteve abaixo do esperado. Atualmente, o protocolo de pirosequenciamento foi aprimorado e permite a obtenção de sequências com tamanho médio de 700-800bp (Dra. Layla F. Martins, CATG/IQUSP, comunicação pessoal).

Novas experiências informáticas de montagem foram realizadas utilizando-se os programas MIRA (http://chevreux.org/projects_mira.html) e Cross_Match (<http://www.phrap.org/phredphrap/general.html>) juntamente com cuidadosas

intervenções manuais, resultando na versão 4.0 que consideramos como final do genoma da cepa J1a12 para efeito das análises que serão discutidas adiante. A figura 6 mostra o mapeamento dos “reads” na sequência consenso do cromossomo principal, demonstrando uma distribuição razoavelmente uniforme ao longo de todo o genoma. Os picos correspondem a regiões que contêm um maior número de *reads* mapeados nela. Esta montagem está sendo validada tanto por análises informáticas adicionais e como experimentalmente.

Figura 6: Mapeamento dos *reads* na sequência consenso do cromossomo principal da cepa J1a12. As sequências de todos os *reads* gerados no pirosequenciamento (*shotgun* e PET) da cepa J1a12 foram mapeadas na sequência consenso do cromossomo principal (versão 4.0) com o programa Cross-Match e sua distribuição visualizada com a ferramenta gráfica do programa Consed.

A montagem final (versão 4.0) resultou em um cromossomo principal de 2.788.789 pb e dois plasmídeos, pXF51 com 51.180 pb e o pXF27 com 27.268 pb e foi então submetida para anotação automática na plataforma IMG/ER (*Integrated Microbial Genomes-Expert Review*, <http://img.jgi.doe.gov/>). A Tabela 4 resume as características gerais do genoma de J1a12.

Tabela 4. Características gerais do genoma da cepa J1a12

Características	Cromossomo	pXF51	pXF27
Tamanho	2.788.789 pb	51.180 pb	27.268 pb
G+C (%)	52.9%	49.6%	49.4%
Genes codificadores de proteínas	2779	73	38
Genes codificadores de proteínas com função predita	1822 (64.13%)	37 (50.68%)	20 (52.63%)
Genes de rRNA	6	-	-
rRNA 5S	2	-	-
rRNA 16S	2	-	-
rRNA 23S	2	-	-
tRNA genes	51	-	-
Outros genes de RNA	7	-	-
tmRNA	1		
SNP-RNA	1		
RNase P-RNA	1		
intron Grupo II -RNA	1		
6S/SsrS RNA	1		
yybP-ykoY líder	1		
TPP <i>riboswitch</i>	1		

O genoma da cepa J1a12 é aparentemente maior (2.788.789 pb) que o genoma da cepa de referência 9a5c (2.679.305 pb), apesar das deleções e genes ausentes previamente identificados (Koide *et al.*, 2004). Estes dados indicam que o genoma da J1a12 é o maior genoma de *X. fastidiosa* entre os já sequenciados.

Além disso, o sequenciamento revelou a existência de dois plasmídeos na cepa J1a12, sendo que um deles (pXF51) é praticamente idêntico ao plasmídeo pXF51 (Marques *et al.*, 2001) encontrado na cepa 9a5c (51.180 pb na J1a12 versus 51.158 pb na cepa 9a5c) e um novo plasmídeo de 27.268 bp, que denominamos pXF27. Este plasmídeo apresenta extensa similaridade com o plasmídeo pVEIS01 de *Verminephrobacter eiseniae* (Pinel *et al.*, 2008) e com plasmídeos de cepas de *X. fastidiosa* que infectam amoreira (Lee *et al.*, 2010; Stenger *et al.*, 2010) e videira (Schreiber *et al.*, 2010). A análise destes plasmídeos e a verificação experimental da sua presença em J1a12 e em outros isolados de *X. fastidiosa* serão apresentadas adiante (item 4.4).

Os genomas de J1a12 e 9a5c apresentam características gerais conservadas. Por exemplo, os genes de rRNA 16S, 23S e 5S também estão organizados em dois operons na cepa J1a12 e apresentam 100% identidade com os seus correspondentes na cepa 9a5c. A anotação realizada na plataforma IMG/ER apontou a presença de 51 tRNAs codificados no genoma da cepa J1a12, destes 49 são idênticos a 49 dos 50 tRNAs preditos no genoma da cepa 9a5c. Um dos tRNAs preditos em 9a5c e que não tem correspondente em J1a12 é um pseudogene. Os dois tRNAs adicionais encontrados no genoma da J1a12 correspondem a cópias dos tRNA-Lys-TTT e tRNA-Ser-CGA.

O cromossomo principal de J1a12 também contém os 7 genes de RNAs de outras categorias preditos na anotação do genoma de 9a5c (Tabela 4).

A comparação entre as sequências de nucleotídeos do cromossomo principal de J1a12 e 9a5c indica a ocorrência de inversões e rearranjos de várias regiões genômicas, como esquematiza o diagrama mostrado na Figura 7. Estas quebras na colinearidade entre as sequências de ambos os genomas ocorre principalmente em regiões de sequências de prófagos, assim como já observado na análise comparativa do genoma da cepa Temecula1 com a 9a5c (Van Sluys *et al.*, 2003).

O genoma da J1a12 exibe 12 regiões associadas a fagos, incluindo 7 regiões já identificadas no genoma de 9a5c. As regiões adicionais de fagos respondem pelo menos por parte da sequência adicional de 109.484 pb verificada no genoma da J1a12. Na Tabela 5 estão listadas algumas dessas regiões maiores, sendo que as

outras são regiões muito pequenas que se repetem ao longo do genoma. A figura 8 esquematiza a organização das regiões de sequências relacionadas a fagos nos genomas da J1a12 e 9a5c. Interessantemente, 5 novas regiões relacionadas a fagos em J1a12, ausentes em 9a5c, mas que apresentam similaridade com regiões de fagos encontrados nas cepas 3124, Fb7, Hib4, Pr8x e Dixon. Estas novas regiões de fagos incluem genes tais como XFJ1_00012120, XFJ1_00017440/XFJ1_00017450, XFJ1_00023200, XFJ1_00009000/XFJ1_00009010 que codificam, respectivamente, toxina Zonula occludens (Zot), um putativo módulo toxina/antitoxina, uma subunidade de uma protease Clp, e um módulo toxina/antitoxina RelE/RelB, além de várias outras proteínas hipotéticas.

As sequências adicionais do cromossomo principal J1a12 (total de 109.484 pb), relativamente ao cromossomo de 9a5c, correspondem a regiões exclusivas de J1a12 e também a duplicações. Essas regiões estão distribuídas em 8 regiões codificando predominantemente genes de proteínas relacionadas a fagos e outras hipotéticas. A Tabela 5 apresenta as coordenadas e algumas características das maiores regiões; o restante são regiões muito pequenas e estão espalhadas ao longo do genoma (dados não mostrados).

Figura 7. Alinhamento do cromossomo principal das cepas 9a5c e J1a12. O diagrama do tipo “Dotplot” foi gerado a partir da comparação da sequência de nucleotídeos dos dois genomas com o programa MUMmer (Delcher *et al.*, 2003) disponível na plataforma IMG/ER. As regiões de similaridade encontradas na mesma orientação estão indicadas em azul e em vermelho, estão indicadas regiões de similaridade em orientação oposta (inversão). As setas indicam as coordenadas dos alinhamentos para alguns pontos selecionados como exemplo.

Figura 8. Organização das regiões genômicas relacionadas a fagos no cromossomo principal das cepas J1a12 e 9a5c. Regiões relacionadas a fagos estão indicadas pela sigla xfp e numeradas de acordo com (de Mello Varani *et al.*, 2008). As ilhas genômicas (gi) e regiões típicas de cepas CVC (cvc-r) estão indicadas. Regiões de fagos encontradas no genoma de J1a12 mas ausentes em 9a5c estão indicadas pela sigla xfpJ. Blocos verdes e vermelhos representam respectivamente, fagos completos e fragmentos. Os pontos pretos no diagrama marcam as posições a cada 0,5 Mpb.

Tabela 5. Coordenadas das regiões adicionais de J1a12

Coordenadas genômicas	Tamanho (pb)	Características principais
1939127...1941300	2174	Contém somente genes para proteínas hipotéticas
2219693...2221652	1960	Contém somente genes para proteínas hipotéticas; pertence ao fago xfpJ2; parte da sequência presente em regiões de outros genomas
1720822...1736964	16143	Contém proteínas hipotéticas e relacionadas a fagos; pertence ao fago xfpJ1; parte da sequência presente em regiões de outros genomas
1955444...1986882	31439	J1a12 apresenta outra região de 13000 pb semelhante a essa; contém subunidade de protease Clp, DNA primase, DNA metilase, hipotéticas e de fagos
821427...825570	4144	Contém proteínas hipotéticas e de fagos; parte da sequência presente em regiões de outros genomas
486100...524351	38.252	Contém proteínas relacionadas a fagos, proteínas hipotéticas, subunidade de protease Clp, DNA primase; parte da sequência presente em regiões de outros genomas

4.3.2. Análises comparativas do genoma de J1a12

As análises comparativas das primeiras versões do genoma de J1a12 foram realizadas utilizando a anotação automática feita no servidor RAST (*Rapid Annotation using Subsystem Technology*) (Aziz *et al.*, 2008), as quais orientaram experimentos para validação de algumas diferenças genômicas observadas e que

foram sendo realizados paralelamente à obtenção da versão final da sequência completa do genoma da J1a12, como por exemplo, a existência dos plasmídeos pXF51 e pXF27 (ver item 4.4).

A versão mais recente da montagem (V4.0) do genoma de J1a12 foi submetida à anotação automática na plataforma IMG/ER (Markowitz *et al.*, 2012). Esta plataforma utiliza um processamento de anotação bastante completo e oferece boa variedade de ferramentas para comparação de genes e genomas. O resumo da comparação do genoma da cepa J1a12 com as cepas 9a5c e Temecula está apresentado na Tabela 6.

Tabela 6. Análises comparativas do genoma de J1a12

Cepa		J1a12	9a5c	Temecula
Características				
Tamanho do Genoma %GC				
	Cromossomo	2.788.789 pb	2.679.306 pb	2.519.802 pb
		52,9%	52,7%	51,8%
	pXF51	51.180 pb	51.158pb	---
		49.6%	50%	---
	pXF27	27.268 pb	---	---
		49,4 %	---	---
	pXF1.3	---	1.285 pb	1.345 pb
		---	56%	54%
Genes codificadores de proteínas (CDS)^a				
	Cromossomo	2.779	2.765 (2470 ^b)	2.066
	pXF51	73	65	---
	pXF27	38	---	---
	pXF1.3	---	2	2
% CDS com função predita		~64%	44% (68% ^b)	~61%
CDS de J1a12 com ortólogos nas cepas indicadas				
≥80% identidade/Valor E ≤10 ⁻¹⁰		---	2.367 (85,1%)	1.993 (71,7%)
≥60% identidade/Valor E ≤10 ⁻¹⁰		---	2.406 (86,5%)	2.060 (74,1%)
≥30% identidade/Valor E ≤10 ⁻¹⁰		---	2.425 (87,2%)	2.106 (75,7%)

^aO número de genes pode variar em função do critério de anotação utilizado em cada genoma e da época em que a anotação foi realizada

^bApós a reanotação do genoma de 9a5c (<http://www.xylella.lncc.br/>)

Apesar da variação no número total de CDS anotadas nos genomas das cepas 9a5c e Temecula e que reflete critérios de anotação utilizados em cada caso, nossas análises mostraram que 85-87% das CDS anotadas automaticamente no genoma de J1a12 tem ortólogos em 9a5c. Como esperado, a porcentagem é menor (72-75%) em relação à cepa Temecula.

Verificamos que 14 dos 15 genes preditos em 9a5c com mutações pontuais que geram *frameshift* (Simpson *et al.*, 2000) apresentam alterações semelhantes em J1a12. Destacamos os genes das subunidades do sistema de restrição do tipo I (XF0295 e XF0297), da transcriptase reversa (XF1775) vizinha ao intron catalítico do grupo II e de uma proteína associada a reparo de DNA (XF1806) além de XF2466 (precursor da poligalacturonase).

Já foi verificado que, além de 9a5c, outras cepas isoladas de citros e de cafeeiros apresentam uma mutação no gene da poligalacturonase que supostamente inviabiliza a expressão de uma enzima funcional, o que tem sido associado a um fenótipo menos agressivo destas cepas em relação às cepas que infectam videiras (Van Sluys *et al.*, 2003; Chatterjee *et al.*, 2008).

Interessantemente, um dos genes (XF1812) originalmente indicados como portadores de mutações de *frameshift* na cepa 9a5c foi, aparentemente, incorretamente anotado. Este locus contém uma ORF com domínio de cisteína-proteases com peptídeo sinal para secreção, além de estar presente no genoma das demais cepas de *X. fastidiosa* já sequenciadas. Além disso, em J1a12 o locus da XF1241 (que em 9a5c possui uma mutação de *frameshift*) apresenta uma alteração que restaura a sequência codificadora da enzima aconitase A (ver tabela 8 adiante).

Por outro lado, 342 e 655 CDS anotadas pela plataforma IMG no genoma de J1a12 não tem, respectivamente, ortólogos anotados nas cepas 9a5c e Temecula. Entretanto, a maioria destas ORFs, respectivamente 87% e 70%, são hipotéticas ou refletem erros da anotação automática, tais como fragmentos de ORFs decorrentes de mudanças de fase causadas por INDELs, muitas destas relacionadas a pseudogenes. Em outros casos, a CDS, embora codificada no genoma, não foi previamente anotada no genoma das cepas 9a5c ou Temecula. Ao final destas análises comparativas tanto das sequências de nucleotídeos como de aminoácidos do conjunto de 342 CDS, confirmamos que 59 CDS anotadas em J1a12 não têm ortólogos na cepa 9a5c, a maioria CDS hipotéticas. Na Tabela 7 destacamos algumas destas CDS indicando, se for o caso, seus ortólogos na cepa Temecula. As regiões genômicas correspondentes a algumas destas CDS estão esquematizadas na Figura 9.

Interessantemente outras CDS presentes no genoma de 9a5c codificam proteínas supostamente equivalentes a 3 das CDS de J1a12 sem ortólogos em 9a5c e listadas na Tabela 7. Por exemplo, o genoma de 9a5c codifica uma aminotransferase (XF2396), duas recombinases da família de integrases de fagos (XF1425 e XF1483) e duas CDS com similaridade parcial à toxina de zonula

occludens (Zot) (XF1870 e XF1875) com ortólogos tanto em J1a12 como em Temecula. A família Zot compreende proteínas bacterianas e virais similares a uma enterotoxina (Zot) produzida por bacteriófagos de *Vibrio cholerae* capaz de alterar a permeabilidade do epitélio intestinal. Proteínas similares a Zot estão descritas em bacteriófagos que infectam outras bactérias patogênicas como *Stenotrophomonas maltophilia* e *Rastonia solanacearum* (Hagemann *et al.*, 2006; Murugaiyan *et al.*, 2011), embora não existam evidências de seu papel na patogenicidade.

Tabela 7. CDS anotadas no genoma de J1a12 sem ortólogos em 9a5c

CDS \ Cepa	J1a12	9a5c	Temecula
Aspartate/tyrosine/aromatic aminotransferase	XFJ1_00006470	Ausente	PD1474
Predicted amidohydrolase	XFJ1_00006480	Ausente	PD1473
Uncharacterized conserved small protein	XFJ1_00009570	Ausente	Ausente
Bacteriophage coat protein B	XFJ1_00012090	Ausente	Ausente
Similar a toxina de zonula occludens (Zot)	XFJ1_00012120	Ausente	PD0915
Recombinase sítio-específica – integrase de fago (XerD)	XFJ1_00012140	Ausente	Ausente
Regulators of stationary/sporulation gene expression	XFJ1_00022920	Ausente	Ausente
Predicted nucleic-acid-binding protein, contains PIN domain	XFJ1_00022930	Ausente	Ausente

Figura 9. Esquema de duas regiões genômicas selecionadas indicando CDS de J1a12 ausentes em 9a5c ou Temecula. As CDS listadas na Tabela 6 estão mostradas no esquema. Note que a recombinase XerD está adjacente ao gene do tRNA-Ser-CGA, que corresponde a borda da região do fago XfpJ5 (Figura 8).

Análises comparativas adicionais com as primeiras versões do genoma de J1a12 no servidor RAST (Aziz *et al.*, 2008) foram realizadas para buscas de ortólogos de ~500 CDS da cepa 9a5c na sequência genômica de J1a12 anotada automaticamente. Utilizando o programa BLAST disponível no próprio servidor RAST, as sequências de nucleotídeos dos ortólogos identificados eram recuperadas e os alinhamentos manualmente inspecionados para identificação de eventuais inserções/deleções (INDEL) ou polimorfismos. As regiões divergentes identificadas na sequência genômica de J1a12 foram inspecionadas diretamente na montagem, com o auxílio do programa CONSED de modo a verificar a cobertura e número de *reads* naquela região para dar ou não suporte às diferenças observadas.

A análise de INDELs é importante na comparação de genomas, uma vez que contribuem para a divergência de sequências de DNA proximamente relacionadas. Além disso, o significado fisiológico de INDELs que originam mutações de “frameshift” (INDELs internos) e transcritos alterados (INDELs externos) são de particular interesse para identificação de características específicas de linhagens distintas de *X. fastidiosa* nas quais a variação genômica não é claramente evidente em nível macro (Doddapaneni *et al.*, 2007). Por outro lado, inserções ou deleções de nucleotídeos em regiões homopoliméricas longas têm, entretanto, que ser analisadas com cautela visto que podem refletir artefatos do pirossequenciamento (Huse *et al.*, 2007).

Entre as CDS inspecionadas incluímos aquelas potencialmente associadas a processos de virulência e patogenicidade em *X. fastidiosa*, tais como adesão, motilidade, sinalização mediada por di-GMP cíclico, além de CDS identificadas como divergentes em J1a12 em experimentos de hibridização de microarranjos de DNA contendo segmentos de todas as CDS do genoma da cepa 9a5c (Koide *et al.*, 2004). As diferenças observadas foram confirmadas na versão 4.0 da montagem do genoma anotado na plataforma IMG/ER e estão listadas na Tabela 8. Entre as ~500 CDS analisadas, observamos que as sequências de 22 estão completamente ausentes no genoma de J1a12 e pelo menos 20 apresentam polimorfismos, deleções e/ou inserções que geram códons de parada prematuros ou alterações na sequência de aminoácidos predita. Vale destacar que as CDS ausentes ou divergentes em J1a12 descritas na comparação genômica utilizando microarranjos de DNA (Koide *et al.*, 2004) foram confirmadas como registrado na Tabela 8 (CDS sublinhadas). Muitas destas estão relacionadas a regiões de fagos que divergiram ou não estão presentes no genoma de J1a12.

Entre as CDSs ausentes que codificam produtos com função predita em 9a5c estão uma das três cópias de colicinas V (XF0264); uma das três cópias da adesina da fimbria/pilus do tipo I (XF0077); dois genes relacionados à função plasmidial, (XF1589 e XF1590) que estão integrados no cromossomo principal da cepa 9a5c; XF1646 (N-aciltransferase); XF1668, que codifica uma proteína de função desconhecida relacionada a um membro de um sistema toxina-antitoxina, denominada HicB, a qual apresenta um domínio de RNAase, além de XF1250 que codifica arginina deaminase.

A ausência de uma cópia da adesina da fimbria (XF0077) parece contribuir para o fenótipo menos agregativo exibido pela J1a12 no cultivo *in vitro* (Figura 2). Como mostrado na Figura 10, a cepa 9a5c têm 3 cópias deste gene enquanto a Temecula possui apenas 1 cópia e, assim como a J1a12, também exibe um fenótipo menos agregado.

Ortólogos de CDS com funções de adesão apresentam deleções que comprometeriam a expressão da proteína íntegra em J1a12 tais como as três cópias para proteínas do tipo das hemaglutininas (XF0889, XF2196 e XF2775). Contudo não podemos excluir a ocorrência de erros de montagem na sequência destas 3 CDS (~10.000 bp com muitas repetições). Ensaios de PCR e sequenciamento dos produtos amplificados deverão ser realizados para verificarmos a montagem destas regiões genômicas.

O significado funcional da ausência completa do gene que codifica arginina deaminase (XF1250) em J1a12 permanece uma incógnita. O esquema na Figura 10 mostra que esta região sofreu um rearranjo que resultou na deleção da ortóloga de XF1250 em J1a12. Especulamos se a CDS XFJ1_00006480 (amido-hidrolase) ortóloga de PD1473 e ausente em 9a5c (Tabela 7) poderia compensar a ausência da arginina deaminase que também hidrolisa ligações carbono-nitrogênio não peptídicas.

Outras CDS completamente ausentes ou com extensas deleções em J1a12 estão em regiões de fagos e codificam, em sua maioria, proteínas sem função conhecida (XF0496, XF0497, XF0500, XF0501, XF1588-XF1590, XF1646, XF1664, XF1707, XF1708, XF1756, XF1758). Uma CDS considerada ausente pela anotação do IMG/ER, XF1267, que codifica uma 1,4-beta-cellobiosidase foi observada presente neste genoma e, portanto, foi anotada manualmente (Figura 10).

Detectamos a inserção ou deleção de 1 nt que resulta na interrupção prematura da ORF nos ortólogos de XF0156, XF0478, XF1483, XF2059 e XF2260.

Contudo, como são alterações em regiões homopoliméricas podem refletir artefatos do pirosequenciamento (Huse *et al.*, 2007) e não serão discutidas neste trabalho.

Em J1a12, duas cópias de CDS que codificam serina proteases da família dos autotransportadores AT-1 estão aparentemente inativadas (XF0267 e XF1851) (Figura 10). A terceira cópia ortóloga de XF1026 apresenta polimorfismos, mas está preservada em J1a12 (Tabela 8). O papel destas proteases na virulência de *X. fastidiosa* ainda não está esclarecido, embora há relatos de que o nocaute destes genes na cepa Temecula reduza sua virulência (Igo *et al.*, 2010).

A análise comparativa do genoma de J1a12 também revelou que esta cepa parece ser deficiente na expressão de seu arsenal de toxinas protéicas. Como mostrado na Figura 10 e Tabela 8, J1a12 dispõe de 2 das 3 cópias de microcinas da família das colicinas V e apresenta alterações nas três CDS que codificam proteínas similares a hemolisinas (XF1011, XF2407 e XF2759). Além disso, observamos alteração no ortólogo de XF2093 anotada como *precursor of drug resistance protein* da família MFP (membrane fusion protein). Em bactérias Gram-negativas MFPs atuam como proteínas adaptadoras conectando um transportador na membrana interna com uma porina ou um canal da membrana externa (Xu *et al.*, 2011). Verificamos que no genoma de 9a5c, outras 3 CDS que codificam MFPs estão associadas a transportadores assim como XF2093 que forma, aparentemente, um operon com XF2094 (*multidrug-efflux transporter*) que pode estar envolvido no efluxo de moléculas pequenas e peptídeos.

Outra CDS cuja inativação pode estar associada à menor virulência da cepa J1a12 é XF0287 (*regulator of pathogenicity factors B*, *rpfB*). RpfB está envolvido no processamento de DSF (*Diffusible Signaling Factor*) que participa do sistema de sinalização célula-célula dependente da densidade celular (*quorum sensing*). Na cepa Temecula, o nocaute de *rpfB*, altera o perfil de moléculas desta família sintetizadas pela bactéria bem como diminui sua eficiência na colonização e transmissão pelo inseto vetor (Almeida *et al.*, 2011). Desta forma, se esta alteração for confirmada, é provável que a cepa J1a12 apresente deficiências nesta via de sinalização.

Tabela 8. CDS de 9a5c divergentes no cromossomo de J1a12

CDS em 9a5c	Produto predito	Característica em J1a12
<u>XF0075</u>	<i>hypothetical protein</i>	Polimorfismos
<u>XF0077</u>	<i>fimbrial adhesin precursor</i>	Ausente
<u>XF0078</u>	<i>fimbrial adhesin precursor</i>	Polimorfismos

XF0083	<i>fimbrial subunit precursor</i>	Polimorfismos
XF0133	<i>copper resistance protein B precursor</i>	Inserção de 10 aminoácidos em fase
XF0156	<i>cysteine protease</i>	Deleção de 1 nt*/códon de parada prematuro
<u>XF0157</u>	<i>hypothetical protein</i>	Polimorfismos
<u>XF0262</u>	<i>colicin V precursor</i>	Polimorfismos
<u>XF0263</u>	<i>colicin V precursor</i>	Polimorfismos
XF0264	<i>colicin V precursor</i>	Ausente
XF0267	<i>serine protease, autotransporter</i>	Polimorfismos, inserções e deleções/ códon de parada prematuro
XF0287	<i>regulator of pathogenicity factors (rpfB)</i>	Inserção de 20 nt/códon de parada prematuro
XF0373	<i>fimbrial assembly protein (pilQ)</i>	Deleção de 1 nt* /códon de parada prematuro
XF0478	<i>tip-associated fimbrial adhesin (pilY1)</i>	Deleção de 1 nt* /códon de parada prematuro
<u>XF0496</u>	<i>conserved hypothetical protein</i>	Ausente
<u>XF0497</u>	<i>conserved hypothetical protein</i>	Ausente
<u>XF0500</u>	<i>phage-related repressor protein</i>	Ausente
<u>XF0501</u>	<i>conserved hypothetical protein</i>	Ausente
<u>XF0626</u>	<i>hypothetical protein</i>	Ausente
<u>XF0663</u>	<i>hypothetical protein</i>	Região com deleção de ~1000pb/CDS ausentes ou truncadas
<u>XF0665</u>	<i>hypothetical protein</i>	
<u>XF0666</u>	<i>hypothetical protein</i>	
<u>XF0667</u>	<i>hypothetical protein</i>	
<u>XF0668</u>	<i>hemolysin-type calcium binding protein</i>	Deleções em fase/região com repetições
XF0818	<i>endo-1,4-beta-glucanase Cellulase (glycosyl hydrolase family 5)</i>	
XF0889	<i>hemagglutinin-like secreted protein (pspA)</i>	Deleções/várias CDS menores/região com repetições
XF1011	<i>hemolysin-type calcium binding protein</i>	Deleções/CDS menor tamanho
XF1024	<i>outer membrane protein H.8 precursor</i>	Inserção de 10 nt/altera sequência de aminoácidos
XF1026	<i>serine protease</i>	Polimorfismos
XF1241	<i>2-methylisocitrate dehydratase, Fe/S-dependent (aconitase A)</i>	Deleção de 1nt*/restaura a fase de leitura de gene com <i>frameshift</i> em 9a5c e Temecula (XFJ1_00011960)
<u>XF1250</u>	<i>arginine deaminase (rocF)</i>	Ausente
XF1267	<i>1,4-beta-cellobiosidase</i>	Inserção em fase/região com repetições
<u>XF1306</u>	<i>conserved hypothetical protein</i>	Inserção/CDS ausente
XF1483	<i>site-specific recombinase (XerC)</i>	Inserção de 1nt*/códon de parada prematuro
<u>XF1588</u>	<i>hypothetical protein</i>	Ausente
<u>XF1589</u>	<i>plasmid stabilization protein</i>	Ausente

<u>XF1590</u>	<i>plasmid stabilization protein</i>	Ausente
<u>XF1646</u>	<i>UDP-3-O-(R-3-hydroxymyristoyl)-glucosamine N-acyltransferase</i>	Ausente
<u>XF1664</u>	<i>hypothetical protein</i>	Deleções
<u>XF1668</u>	<i>HicB-related protein</i>	Ausente
<u>XF1669</u>	<i>phage-related lysozyme</i>	Polimorfismos
<u>XF1707</u>	<i>hypothetical protein</i>	Ausente
<u>XF1708</u>	<i>hypothetical protein</i>	Ausente
<u>XF1756</u>	<i>hypothetical protein</i>	Ausente
<u>XF1758</u>	<i>hypothetical protein</i>	Polimorfismos
<u>XF1774</u>	<i>DNA methyltransferase</i>	Deleções/CDS ausente
<u>XF1844</u>	<i>ammonium transporter</i>	Deleções/diminui N-terminal da CDS
<u>XF1851</u>	<i>serine protease, autotransporter</i>	Deleções/ altera N-terminal da CDS e elimina peptídeo sinal
<u>XF1860</u>	<i>conserved hypothetical protein</i>	Polimorfismos
<u>XF1862</u>	<i>hypothetical protein</i>	Ausente
<u>XF1873</u>	<i>conserved hypothetical protein</i>	Ausente
<u>XF1874</u>	<i>hypothetical protein</i>	Ausente
<u>XF1877</u>	<i>hypothetical protein</i>	Ausente
<u>XF1878</u>	<i>hypothetical protein</i>	Ausente
<u>XF1884</u>	<i>hypothetical protein</i>	Ausente
<u>XF1968</u>	<i>methyltransferase, type III restriction-modification system</i>	Polimorfismos
<u>XF2046</u>	<i>conjugal transfer protein (trbL)</i>	Inserções/códon de parada prematuro
<u>XF2059</u>	<i>conjugal transfer protein (traE)</i>	Deleção de 1 nt*/códon de parada prematuro
<u>XF2061</u>	<i>antirestriction protein</i>	Deleção de ~340nt
<u>XF2093</u>	<i>precursor of drug resistance protein</i>	Inserção de 4 nt/códon de parada prematuro
<u>XF2185</u>	<i>rare lipoprotein A</i>	Polimorfismos e inserção 6 nt em fase
<u>XF2193</u>	<i>hypothetical protein</i>	Região com deleção 500 bp, inserções e polimorfismos/CDS ausentes ou truncadas
<u>XF2194</u>	<i>hypothetical protein</i>	
<u>XF2195</u>	<i>hypothetical protein</i>	
<u>XF2196</u>	<i>hemagglutinin-like secreted protein</i>	Deleções/CDS íntegra ausente
<u>XF2260</u>	<i>alanyl dipeptidyl peptidase</i>	Inserção de 1nt*/códon de parada prematuro
<u>XF2276</u>	<i>peptide synthase</i>	Inserção de 3 nt em fase
<u>XF2406</u>	<i>hypothetical protein</i>	Polimorfismos
<u>XF2407</u>	<i>bacteriocin/ hemolysin</i>	Polimorfismos e deleções reduzem tamanho da proteína predita de 2407 para 1377 aminoácidos
<u>XF2455</u>	<i>heme ABC transporter ATP-binding protein (ccmA)</i>	Inserção de 156 nt/CDS (XFJ1_00006460) com N-

		terminal mais longo
<u>XF2542</u>	<i>fimbrial protein (Type-4 pilus assembly protein)</i>	Polimorfismos
XF2722	<i>type I restriction-modification system specificity determinant</i>	Inserções/deleções e polimorfismos
XF2723	<i>type I restriction-modification system DNA methylase</i>	Estas duas sequências estão fusionadas em J1a12 (XFJ1_00027770) assim como na cepa Temecula (PD2072)
XF2724	<i>type I restriction-modification system</i>	
XF2725	<i>type I restriction-modification system endonuclease</i>	Inserções e polimorfismos/aumentam tamanho da proteína predita/N-terminal 10 aa mais longo
<u>XF2726</u>	<i>type I restriction-modification system specificity determinant</i>	Sequência de nucleotídeos muito divergente porém codifica CDS ortóloga (XFJ1_00027800)
XF2728	<i>type I restriction-modification system DNA methylase</i>	Polimorfismos
XF2759	<i>hemolysin-type calcium binding protein</i>	Inserções e polimorfismos/CDS com C-terminal mais curto
<u>XF2768</u>	<i>hypothetical protein</i>	Região com deleções/CDS ausentes ou truncadas
<u>XF2770</u>	<i>hypothetical protein</i>	
<u>XF2772</u>	<i>hypothetical protein</i>	
XF2775	<i>hemagglutinin-like secreted protein</i>	Deleções/CDS íntegra ausente

As CDS sublinhadas foram previamente identificadas como ausentes ou divergentes na comparação entre os genomas de J1a12 e 9a5c através de microarranjos de DNA (Koide *et al.*, 2004). O asterisco indica que a inserção ou deleção ocorreu em sequência homopolimérica.

Figura 10. Comparação de regiões genômicas selecionadas de J1a12, 9a5c e Temecula. Representação gráfica de regiões genômicas de J1a12, 9a5c e Temecula para algumas das CDS listadas na Tabela 7. As imagens foram extraídas da plataforma de anotação automática IMG/ER. O símbolo pontilhado da CDS XF1267 em J1a12 foi inserido manualmente, após identificarmos que a ORF está presente, mas não foi reconhecida na anotação automática.

Entre as alterações que nos parecem relevantes na tentativa de explicar características fenotípicas da cepa J1a12, estão alterações em genes relacionados ao sistema de restrição/modificação de DNA (XF1968 e XF2722 a XF2726). O genoma desta cepa assim como o da cepa 9a5c codifica para sistemas de restrição/modificação do tipo I e do tipo III, os quais também estão presentes no

genoma da cepa Temecula, sendo que nesta última também existe o sistema do tipo II, aparentemente exclusivo de *X. fastidiosa subsp. fastidiosa* (Simpson *et al.*, 2000; Van Sluys *et al.*, 2003; Matsumoto e Igo, 2010). Estes sistemas de restrição/modificação protegem as bactérias de DNAs invasores (ex: bacteriófagos e plasmídeos), porém interferem na aquisição de DNA exógeno afetando a troca de material genético interespecies, ou mesmo intercepas.

O sistema de restrição/modificação do tipo I reconhece sequências específicas no DNA alvo, porém a clivagem é realizada muito distante (até 1000bp) do sítio de reconhecimento, diferentemente dos sistemas do tipo II e III. O complexo multimérico do sistema do tipo I é composto pela subunidade de especificidade (Subunidade-S) que é responsável por reconhecer a sequência específica de DNA e por subunidades com atividade de metilase de DNA (Subunidade M) e com atividade de endonuclease de restrição (Subunidade R), que respectivamente metilam o DNA bacteriano protegendo-o da ação da endonuclease. Enquanto as subunidades M e R são conservadas dentro de distintas subfamílias, as subunidades S são variáveis na região que reconhece a sequência alvo (Tock e Dryden, 2005)

O genoma da cepa Temecula possui dois *clusters* de genes do sistema de restrição/modificação do tipo I enquanto que os genomas das cepas 9a5c e J1a12 possuem quatro. Aparentemente, em 9a5c somente dois clusters são funcionais enquanto que em J1a12 três clusters seriam funcionais, pois em um deles a sequência da metilase (XF2723/XF2724) está intacta e não dividida como em 9a5c (Figura 11). Além disso, a comparação das sequências de aminoácidos das subunidades S codificadas nas duas cepas revela diferenças decorrentes dos polimorfismos em suas sequências de nucleotídeos. Uma possibilidade a ser considerada, mas que nos parece pouco provável, é se estas diferenças têm relação com a capacidade destas cepas serem transformadas naturalmente ou via eletroporação. Enquanto a cepa 9a5c é resistente à transformação por DNA, o que é um impedimento para sua manipulação genética, as cepas J1a12 e Temecula são passíveis de transformação de DNA (Monteiro *et al.*, 2001; Matsumoto e Igo, 2010; Kung e Almeida, 2011).

Outro aspecto que vale ressaltar é que, em J1a12, a CDS que codifica uma provável proteína de anti-restrição encontrada em 9a5c (XF2061) apresenta uma deleção de ~128 aminoácidos em sua porção central (Figura 12). Ortólogos de XF2061 também estão presentes nos genomas das cepas Dixon, Ann-1, EB92.1, 3124, Fb7, U24d e Hib4, mas não nos demais genomas sequenciados, incluindo Temecula e Pr8x. XF2061 foi originalmente anotada como DNA primase, e está em

uma região cromossômica que integrou vários genes do plasmídeo conjugativo pXF51. Não há evidência experimental de que XF2061 é expressa e que poderia desempenhar algum papel como proteína de anti-restrição. Porém interessantemente, sua porção carboxiterminal tem similaridade a proteína TraC (envolvida no processo de conjugação) incluindo um domínio carboxi-terminal de topoisomerase/primase enquanto que sua porção amino-terminal é similar (~50% similaridade) a proteínas de anti-restrição ArdC. Entre os mecanismos empregados por fagos e plasmídeos para evitar o sistema de restrição, está a oclusão temporária de sítios de restrição que pode ocorrer pela associação de proteínas, tais como ArdC, com moléculas do DNA invasor até que estas recebam modificações que as protejam da sistema de restrição do hospedeiro (Belogurov *et al.*, 2000; Tock e Dryden, 2005). Estas observações nos levam a especular se as diferenças na sequência da proteína de anti-restrição em J1a12 (Figura 12) lhe conferem características que tenham facilitado a estabilização dos dois plasmídeos e invasão dos fagos existentes nesta cepa.

Figura 11. Comparação da região genômica de J1a12, 9a5c e Temecula que codifica genes do sistema de restrição/modificação do tipo I. Representação gráfica da região genômica de J1a12, 9a5c e Temecula. As CDS indicadas estão referidas na Tabela 8. As imagens foram extraídas da plataforma de anotação automática IMG/ER.

2510343758	MGVTLLQAQALVVDRCRNYPVAAQLSYYIRETVEELYGKCPEGMERMLGAYVTGPGIHPG	60
XF2061	MGVTLLQAQALVVDRCRNYPVAAQLSYYIRETVEELYGKCPEGMERMLGAYVTGPGIHPG	60
2510343758	*****	
2510343758	RCDIPCGNAKDTASFISTLRHEVIGHFGLNTFTREDKKALLDGISAVREKPFMRSLWRIV	120
XF2061	RCDIPCGNAKDTASFISTLRHEVIGHFGLNTFTREDKKALLDGISAVREKPFMRSLWRIV	120
2510343758	*****	
2510343758	EENYAHESESIAKEEVFALYCEGIEPELVHVTTRYVHEQGEQVFRETCDIRTRLNMIIDLD	180
XF2061	EENYAHESESIAKEEVFALYCEGIEPELVHVTTRYVHEQGEQVFRETCDIRTRLNMIIDLD	180
2510343758	*****	
2510343758	SIVLMVAQGLHDLRLQRFNTELNQQFRQEPYMEPKKPFHQVVAERLIEQLKVGTPAPWQR	240
XF2061	SIVLMVAQGLHDLRLQRFNTELNQQFRQEPYMEPKKPFHQVVAERLIEQLKVGTPAPWQR	240
2510343758	*****	
2510343758	PWEPGRGGLLPMNPSTGKRYKGINAIHLSMQQRTDPRWLTYKQAEITLGAQVRRGEKSTSI	300
XF2061	PWEPGRGGLLPMNPSTGKRYKGINAIHLSMQQRTDPRWLTYKQAEITLGAQVRRGEKSTSI	300
2510343758	*****	
2510343758	QYWKDTEKQIKTDDQGRPVLDVKGQPMEEVRLERPRVFFASVFNAEQIDGLPTLQHNEQ	360
XF2061	QYWKDTEKQIKTDDQGRPVLDVKGQPMEEVRLERPRVFFASVFNAEQIDGLPTLQHNEQ	360
2510343758	*****	
2510343758	TKEQTWDAVERAERILEASAAVIHHGEHNRIFYRSTDSIHLADKSQFPSADNYATLLH	420
XF2061	TKEQTWDAVERAERILEASAAVIHHGEHNRIFYRSTDSIHLADKSQFPSADNYATLLH	420
2510343758	*****	
2510343758	ELGHWGHSRLRLDRDMLHPFGSEGYAKEELRAEIASMILGDELIGHDGQHAAYVGSWI	480
XF2061	ELGHWGHSRLRLDRDMLHPFGSEGYAKEELRAEIASMILGDELIGHDGQHAAYVGSWI	480
2510343758	*****	
2510343758	RVLKDDPLEIFRAADAELIQEYLLAFEQKQVQEQATKQQAIAQNPEPLRQEQAVRLPSR	540
XF2061	RVLKDDPLEIFRAADAELIQEYLLAFEQKQVQEQATKQQAIAQNPEPLRQEQAVRLPSR	540
2510343758	*****	
2510343758	MLADQ-----	545
XF2061	MLADQKVAILLRIAREDPNDPFLGVYKGLDIAFDEAFGLVKGMGEQVREVLPLDWNGRV	600
2510343758	*****	
2510343758	RVESDDSEPNRFVVFVQRQNGEYEWLEICETSAAMDSLIDRLAVINAYAEIHEKQARYEE	660
2510343758	-----KVATALRIARDLKSSFHQETLRVVFSEAFGPVVEGVRDVLDPADWNGR	592
XF2061	QNVQLPSITLADQKVATALRIARDLKSSFHQETLRVVFSEAFGPVVEGVRDVLDPADWNGR	720
2510343758	*****	
2510343758	VRVEPSFVVKNSYMLADVGREPTHWRVVFQCCQNGKYEWLETCTSTAVESLVDRLVVID	652
XF2061	VRVEPSFVVKNSYMLADVGREPTHWRVVFQCCQNGKYEWLETCTSTAVESLVDRLVVID	780
2510343758	*****	
2510343758	AYTKENEHERLTAFARIHEKKARNEKYILAAKTAVADNTVQEQQEQYVSLPSLAISAY	712
XF2061	AYTKENEHERLTAFARIHEKKARNEKYILAAKTAVADNTVQEQQEQYVSLPSLAISAY	840
2510343758	*****	
2510343758	AKANEHEWVAAFVTRTHEKQPRHDKDILAAKTAVADNTVQEQQEQDVALPSITVSEKIWL	772
XF2061	AKANEHEWVAAFVTRTHEKQPRHDKDILAAKTAVADNTVQEQQEQDVALPSITVSEKIWL	900
2510343758	*****	
2510343758	DLPYKQKDAAKHAAGNLADGNCAISWDKEEKRWYARPGADLDKIKQWLPSEAKPTNLAT	832
XF2061	DLPYKQKDAAKHAAGNLADGNCAISWDKEEKRWYARPGADLDKIKQWLPSEAKPTNLAT	960
2510343758	*****	
2510343758	KKTWLAVTFEQRHALKEIGGKLPGNNNAVDWDNAACWYANPGADLDKLPWIAIEVVS	892
XF2061	KKTWLAVTFEQRHALKEIGGKLPGNNNAVDWDNAACWYANPGADLDKLPWIAIEVVS	1020
2510343758	*****	
2510343758	QVPALSPEKEFAETLRSMGCLVSGEHPIMDAKKHRIEVEGDKKGEKAGFYVAYLDGHP	952
XF2061	QVPALSPEKEFAETLRSMGCLVSGEHPIMDAKKHRIEVEGDKKGEKAGFYVAYLDGHP	1080
2510343758	*****	
2510343758	YIKNNRTGIELKWKSKGYSLDPEQKAAMQAEAAKLAARAEQQQEQEAVAQRVVKQMD	1012
XF2061	YIKNNRTGIELKWKSKGYSLDPEQKAAMQAEAAKLAARAEQQQEQEAVAQRVVKQMD	1140
2510343758	*****	
2510343758	LVTVIEPTPYLKAKGITPQAGIYTDSDGQKTYIPAIDSQKQWMTQYIQEDGTRKFAKGG	1072
XF2061	LVTVIEPTPYLKAKGITPQAGIYTDSDGQKTYIPAIDSQKQWMTQYIQEDGTRKFAKGG	1200
2510343758	*****	
2510343758	RKEGCFHPIGGLAAIAAPVIVIGEGYATVAQVSQALGYGTVAAFDAGNLEGVAKALHKT	1132
XF2061	RKEGCFHPIGGLAAIAAPVIVIGEGYATVAQVSQALGYGTVAAFDAGNLEGVAKALHKT	1260
2510343758	*****	
2510343758	FDPKPILIAAGDDQSLPNNPGKTKAQEAARAVDGTAFPFIFAPGEQADNPKGFTDFNDLA	1192
XF2061	FDPKPILIAAGDDQSLPNNPGKTKAQEAARAVDGTAFPFIFAPGEQADNPKGFTDFNDLA	1320
2510343758	*****	
2510343758	TKSALGIEAVERQVHAASKAIDNHQKRVQQQTQQQGQHEQKQERRRLRVR	1243
XF2061	TKSALGIEAVERQVHAASKAIDNHQKRVQQQTQQQGQHEQKQERRRLRVR	1371

Figura 12. Alinhamento da sequência de XF2061 e sua ortóloga em J1a12 (proteína de anti-restrição).

4.4. Caracterização dos plasmídeos de pXF51 e pXF27 da cepa J1a12

Como já mencionado (Tabelas 4 e 6), o sequenciamento do genoma de J1a12 revelou a existência de dois plasmídeos, pXF51, praticamente idêntico ao plasmídeo pXF51 (Marques *et al.*, 2001) encontrado na cepa 9a5c, e pXF27, um novo plasmídeo, que apresenta similaridade com o plasmídeo pVEIS01 de *Verminephrobacter eiseniae* (Pinel *et al.*, 2008) e com plasmídeos de cepas de *X. fastidiosa* que infectam amoreira (Lee *et al.*, 2010; Stenger *et al.*, 2010), videira (Schreiber *et al.*, 2010) e espirradeira (Bhattacharyya *et al.*, 2002).

4.4.1. Análises comparativas de pXF51

A comparação das CDS codificadas no plasmídeo pXF51 das cepas J1a12 e cepa 9a5c mostrou poucas diferenças, as quais aparentemente não alteram as funções preditas neste plasmídeo (Marques *et al.*, 2001). Em J1a12, a CDS ortóloga de XFa0044 possui deleções, e assim como em 9a5c aparentemente não codifica a proteína de conjugação TrbE funcional. Por outro lado, uma cópia aparentemente funcional de TrbE encontra-se integrada no cromossomo destas cepas.

Duas outras CDSs ortólogas em J1a12 de XFa0016 (proteína de conjugação VirD4) e XFa0037 (proteína de conjugação TrbL) apresentam inserções que geram códon de terminação da tradução prematuro.

Entretanto, devido a presença de dois sistemas de conjugação diferentes no pXF51, o sistema VirB-VirD e o sistema *trb*, podemos especular que ocorra sobreposição de funções compensando prováveis deficiências em algum componente destes dois sistemas, mantendo, assim, a capacidade do plasmídeo de ser transferido por conjugação. Além disso, um segmento de pXF51 (~9.100 bp) está integrado no cromossomo de 9a5c (Marques *et al.*, 2001) e também de J1a12, o qual codifica componentes do sistema *trb* aparentemente funcionais.

Polimorfismos em outros genes de pXF51 de J1a12 foram observadas nesta comparação, porém todos codificam CDS hipotéticas.

4.4.2. Análises comparativas de pXF27

A existência do plasmídeo pXF27 em J1a12 foi revelada na análise dos resultados da primeira corrida de pirosequenciamento (da Silva, 2010). Sua montagem foi refinada com os novos dados de pirosequenciamento que obtivemos neste trabalho. Como já mencionado, pXF27 apresenta similaridades com plasmídeos descritos em cepas norte-americanas de *X. fastidiosa* (Bhattacharyya *et al.*, 2002; Lee *et al.*, 2010; Schreiber *et al.*, 2010; Stenger *et al.*, 2010) e com o

plasmídeo pVEIS01 de *V. eiseniae* (Pinel *et al.*, 2008). A figura 13 exemplifica a similaridade entre pXF27 e pXF-RIV11, um plasmídeo isolado de cepas de amoreira.

A presença de pXF27 na cepa J1a12 sugere que ancestrais de uma gama-proteobacteria (*X. fastidiosa*) e uma beta-proteobactéria (*V.eiseniae*) provavelmente trocaram material genético. Ainda não se pode afirmar se essa transferência horizontal gênica ocorreu diretamente, ou indiretamente via intermediários, nem a escala de tempo envolvida. A observação da presença de plasmídeos semelhantes em cepas norte-americanas, tanto da Califórnia como do Texas, tanto em amoreiras como em videiras, sugere que este plasmídeo provavelmente deve ter sido adquirido por um ancestral destas cepas antes da separação geográfica entre elas. Diferentes linhagens de *X. fastidiosa* foram mostradas infectando um único hospedeiro no campo e dessa forma a disseminação deste plasmídeo pode ter ocorrido. A probabilidade de que vários eventos de conjugação ocorram de forma independente é muito menor do que um único evento. Dessa forma, *X. fastidiosa* pode ter tido meios e oportunidades para trocar genes entre linhagens e entre outras espécies por conjugação, incluindo genes que podem ser determinantes da abrangência de hospedeiro e/ou virulência. Entretanto, ainda não é possível afirmar se o sistema de secreção do tipo IV descrito para os plasmídeos de *X. fastidiosa* é funcional para transferência de DNA por conjugação (Lee *et al.*, 2010; Stenger *et al.*, 2010).

A anotação automática apontou a presença de 39 CDS em pXF27 que estão listadas na Tabela 9, algumas delas ortólogas de ORFs previamente identificadas em pXF-RIV11 e/ou em um plasmídeo da cepa Ann-1.

Figura 13. Alinhamento de pXF27 e pXF-RIV11. O diagrama do tipo “Dotplot” em (A) foi gerado a partir da comparação da sequência de nucleotídeos dos dois plasmídeos com o programa MUMmer (Delcher *et al.*, 2003) disponível na plataforma IMG/ER. As regiões de similaridade encontradas na mesma orientação estão indicadas por linhas em vermelho.

Tabela 9. CDS do plasmídeo pXF27 da cepa J1a12 e suas respectivas funções preditas

CDS pXF27	Produto predito	Tamanho predito em aminoácidos	Ortólogo em pXF- RIV11
1	<i>Plasmid encoded RepA protein</i>	367	
2	<i>DNA invertase</i>	147	P25
3	<i>Programmed cell death antitoxin YdcD</i>	72	
4	<i>Programmed cell death toxin YdcE</i>	108	
5	<i>RelB/StbD replicon stabilization protein (antitoxin to RelE/StbE)</i>	76	
6	<i>RelE/StbE replicon stabilization toxin</i>	89	
7	<i>hypothetical protein</i>	54	
8	<i>hypothetical protein</i>	67	P21
9	<i>hypothetical protein</i>	66	
10	<i>TrbC/VIRB2 family.</i>	72	P19
11	<i>Type IV secretory pathway, VirB3 component</i>	127	P18
12	<i>Type IV secretory pathway, VirB4 component</i>	769	P17
13	<i>Type IV secretion system protein (VirB5)</i>	221	P16
14	<i>hypothetical protein</i>	98	
15	<i>Type IV secretory pathway, VirB6 component</i>	327	P11
16	<i>Type IV secretory pathway, component VirB8</i>	155	P10
17	<i>Type IV secretory pathway, VirB9 component</i>	199	P9
18	<i>hypothetical protein</i>	204	
19	<i>hypothetical protein</i>	56	
20	<i>Type IV secretory pathway, VirB10 component</i>	125	P8
21	<i>Type IV secretory pathway, VirD4 component</i>	1105	P6
22	<i>hypothetical protein</i>	275	P5
23	<i>hypothetical protein</i>	101	
24	<i>hypothetical protein</i>	210	
25	<i>hypothetical protein</i>	94	
26	<i>hypothetical protein</i>	115	
27	<i>Transposase</i>	365	
28	<i>Transposase</i>	146	
29	<i>Plasmid maintenance toxin/ Cell growth inhibitor / PemK toxin</i>	120	P3
30	<i>Programmed cell death antitoxin PemI</i>	87	
31	<i>Zeta toxin family protein</i>	366	P1
32	<i>Epsilon antitoxin</i>	84	
33	<i>Bacterial mobilisation protein (MobC).</i>	168	P33
34	<i>hypothetical protein</i>	118	P32
35	<i>CobQ/CobB/MinD/ParA nucleotide binding domain.</i>	250	P31
36	<i>Transcriptional activator TraM.</i>	147	P30
37	<i>Growth inhibitor</i>	108	
38	<i>hypothetical protein</i>	96	

39	<i>hypothetical protein</i>	45
----	-----------------------------	----

Assim como pXF-RIV11, pXF27 codifica todo o aparato do sistema de secreção do tipo IV com exceção das CDS correspondentes a VirB1 e VirB7. Ambas são descritas como sendo dispensáveis na formação do sistema de secreção do tipo IV, sendo que a subunidade VirB1 facilita mas não é exigida para a montagem do aparato e a VirB7 não foi encontrada em alguns sistemas já descritos (Christie *et al.*, 2005). Desta forma podemos sugerir que este plasmídeo seja conjugativo.

Além disso, de acordo com análises utilizando o IMG/ER, pXF27 codifica três prováveis CDS com similaridade a PemK, indicando a possível existência do sistema toxina/antitoxina similar a PemK/PemI. Também codifica uma CDS similar a toxina Zeta, que compõe o sistema toxina-antitoxina Zeta/Êpsilon; a antitoxina épsilon está localizada adjacente à CDS da toxina Zeta. Comparações das sequências em bases de dados públicas sugerem que as CDS 3 e 4 do pXF27 correspondem ao sistema antitoxina/toxina YdcD/YdcE (Gogos *et al.*, 2003) e, as CDS 5 e 6, ao sistema RelB/RelE (Gottfredsen e Gerdes, 1998).

Sistemas bacterianos de toxina-antitoxina (TA) são pequenos sistemas gênicos encontrados em plasmídeos assim como em cromossomos e fazem parte do conteúdo flexível do genoma. Consistem em geral de dois componentes, uma toxina estável e uma antitoxina lábil. Sistemas de TA são de três tipos, dependendo da natureza da antitoxina e modo de ação. Enquanto toxinas são sempre proteínas, antitoxinas podem ser tanto RNAs (tipo I e III) ou proteínas (tipo II). A antitoxina dos sistemas do tipo I é um pequeno RNA (antisense ou adjacente e divergente ao gene da toxina) mostrando complementaridade com o mRNA da toxina, regulando a expressão da toxina pela inibição da tradução da mesma. As toxinas dos sistemas do tipo I são pequenas proteínas hidrofóbicas que causam danos em membranas celulares bacterianas. Em sistemas do tipo II, a antitoxina é uma proteína pequena e instável que sequestra a toxina através da formação de um complexo proteico. Sistemas tipo I não se movem através de transferência horizontal gênica, mas evoluíram de duplicação linhagem específica e, portanto, parecem estar menos disseminados e abundantes quando comparados com o sistema do tipo II. Com isso, muito mais informação está disponível para sistemas do tipo II, especialmente em termos de seus papéis biológicos. Sistemas do tipo II estão organizados em operons, com o gene *upstream* normalmente codificando para a proteína da antitoxina. Nove famílias de toxinas foram definidas até o momento. Normalmente pode-se observar a co-existência entre dois ou mais destes sistemas dentro de um

genoma. Em geral, estes sistemas não interagem, o que pode refletir evolução independente e, portanto contribuir para a sua persistência. A presença de sistemas TA nos cromossomos pode direcionar evolução de seus homólogos codificados em plasmídeos e, os sistemas de plasmídeos que escapam do cromossomo podem ser selecionados. Tais sistemas codificados no cromossomo podem perder suas propriedades aditivas o que pode ser o primeiro sinal de degeneração (Van Melderren e De Bast, 2009; Van Melderren, 2010). Os sistemas TA são responsáveis pela manutenção do plasmídeos nas células-filhas a cada processo de replicação, sendo que as células que não receberem cópias do plasmídeo, não conseguirão repor a quantidade de proteína antitoxina, a qual é mais lábil que a toxina e, conseqüentemente, liberariam a toxina livre para atuar e matar a célula (Van Melderren, 2010).

Estudos realizados com o sistema PemK/PemI de pXF-RIV11 de *X. fastidiosa* demonstraram que este confere maior estabilidade a construções plasmidiais contendo o gene *trfA* (proteína de replicação plasmidial) e a origem de replicação *oriV*. PemK foi validada como uma toxina que inibe o crescimento bacteriano e PemI foi mostrada como sendo a antitoxina que previne a ação de PemK pela direta ligação a ela (Stenger *et al.*, 2010). Posteriormente, PemK teve sua ação descrita com mais detalhes, no que diz respeito ao seu alvo de ação e sítios preferenciais de ligação. Foi mostrado que a atividade de nuclease desta toxina é específica para moléculas de RNA fita simples (ssRNA), aparentemente não atuando sobre dsRNA, dsDNA e ssDNA. A antitoxina PemI foi confirmada como inibidora da atividade de ribonuclease da toxina, sendo sua ligação reversível. Além disso, foi mostrado que PemK reconhece sequências específicas para a clivagem, sendo UAC o sítio de reconhecimento preferencial com eficiência de clivagem variando de acordo com a base seguinte na extremidade 3' a esta sequência (U e G em preferência a C e A). *In vivo*, o alvo provável de PemK pode ser mRNAs (Lee *et al.*, 2010).

Outro elemento de toxina/antitoxina comumente encontrado em plasmídeos e cromossomos de vários patógenos são os sistemas épsilon/zeta (ϵ/ζ). Enquanto que os sistemas codificados no plasmídeo garantem sua manutenção na população, homólogos codificados no cromossomo parecem contribuir para a virulência do patógeno (Mutschler e Meinhardt, 2011). Este sistema foi originalmente descoberto em um plasmídeo de baixo número de cópias, pSM19035, do patógeno Gram-positivo *Streptococcus pyogenes* (Ceglowski *et al.*, 1993). Originalmente, acreditava-se que o sistema ϵ/ζ estava confinado a bactérias Gram-positivas, mas foi

recentemente identificado em plasmídeos IncP1 de *Neisseria gonorrhoeae* (Pachulec e van der Does, 2010).

Trata-se de um sistema codificado por um operon bicistrônico com ϵ/ζ , o qual é regulado pela proteína repressora transcricional Omega (ω). A interação ocorre em um complexo heterotetramérico no qual duas antitoxinas inibem duas toxinas (ϵ_2/ζ_2). Na ausência de síntese continuada de epsilon, a toxina zeta é liberada do complexo toxina/antitoxina a partir da degradação da antitoxina pela ação de proteases, levando à morte celular. Foi demonstrado que toxinas ζ fosforilam especificamente, utilizando ATP e Mg^{2+} , o precursor de peptidoglicano UDP-N-acetilglicosamina (UNAG) no grupo 3'-OH da porção N-acetilglicosamina definindo uma nova família de pequenas quinases. O produto da reação tóxico, UDP-N-acetilglicosamina-3-fosfato (UNAG-3P), se acumula no citosol e MurA, uma enzima essencial que catalisa o primeiro passo da síntese de peptidoglicano, não consegue utilizá-lo para o processo de síntese. A molécula de UNAG e a enzima MurA são essenciais para a formação de vários glicoconjugados, além do peptidoglicano, como lipopolissacarídeos. Portanto, toxinas zeta são as primeiras toxinas de sistemas toxina/antitoxina que têm como alvo direto a síntese de parede celular (Van Melder, 2010; Mutschler *et al.*, 2011; Mutschler e Meinhart, 2011).

A anotação do plasmídeo pXF-RIV11, revelou que o plasmídeo contém homólogos dos elementos de replicação do grupo de incompatibilidade P-1 (*trfA* e *oriV*). A iniciação da replicação em plasmídeos IncP-1 exige a proteína TrfA, codificada no próprio plasmídeo, a qual se associa a sequências na origem de replicação *oriV* (Lee *et al.*, 2010; Stenger *et al.*, 2010). Interessantemente TrfA não está codificada em pXF27, que por sua vez codifica a proteína de replicação plasmidial RepA, que apresenta 100% de identidade com sua ortóloga codificada em um plasmídeo da cepa Ann-1, o que leva a suposição que pXF27 pertença a grupo de incompatibilidade distinto de pXF-RIV11.

Plasmídeos pertencentes ao mesmo grupo de incompatibilidade não podem co-existir na mesma célula. A cepa J1a12 contém dois plasmídeos relativamente grandes (pXF51 e pXF27), sugerindo que eles sejam de grupos Inc diferentes e contenham origens de replicação distintas. O pXF51 não teve sua classificação determinada quando ao seu grupo de incompatibilidade uma vez que não foi possível identificar uma típica origem de replicação, entretanto, foi encontrada uma origem de transferência por conjugação *oriT* de plasmídeos IncP (Marques *et al.*, 2001).

4.4.3. Ensaios para identificação de pXF51 e pXF27 em *X. fastidiosa*

Assim como outras bactérias, *X. fastidiosa* possui uma variedade de plasmídeos, alguns dos quais têm sido caracterizados molecularmente. Como mencionado acima, o sequenciamento do genoma da linhagem 9a5c revelou a presença denominado pXF51. Qin e Hartung (2001) caracterizaram um plasmídeo de cerca de 6 kpb (pXF5823) a partir de uma linhagem também isolada de citros. Outras linhagens de *X. fastidiosa* como por exemplo, as linhagens Ann-1 (isolada de espiroleira) e Dixon (isolada de amendoeira) possuem plasmídeos distintos de cerca de 30 kpb (Bhattacharyya *et al.*, 2002), a linhagem M23, isolada de amendoeira e que causa doença de Pierce em videiras, também possui um plasmídeo de cerca de 38 kpb (Chen *et al.*, 2010) e, como já destacamos, quatro linhagens de amoreira (Stenger *et al.*, 2010) e uma de videira (Schreiber *et al.*, 2010) possuem um plasmídeo de cerca de 25 kpb, semelhantes entre si e com alta similaridade com um plasmídeo de *Verminephrobacter eiseniae*. Além disso, várias linhagens de *X. fastidiosa* possuem um plasmídeo de cerca de 1.3 kpb, como por exemplo, as linhagens 9a5c (Simpson *et al.*, 2000), Temecula1 (Van Sluys *et al.*, 2003) e EB92-1 (Zhang *et al.*, 2011). Estes relatos indicam que a diversidade de elementos extracromossomais associados com *X. fastidiosa* é extensa mas pouco caracterizada no que diz respeito ao conteúdo e função de genes (Stenger *et al.*, 2010).

Com o objetivo de confirmar a presença dos plasmídeos pXF27 e pXF51 revelados através do pirosequenciamento da cepa J1a12, nós investigamos o conteúdo plasmidial desta e de outras cepas utilizadas neste trabalho. Para tal realizamos extrações de DNA plasmidial das cepas 9a5c, J1a12, Pr8x e Hib4 a partir de 50 mL de cultivo em PWG líquido. O rendimento e quantificação do DNA plasmidial purificado está mostrado na Tabela 10.

Tabela 10. Análise quantitativa e qualitativa do DNA plasmidial isolado das cepas J1a12, 9a5c, Pr8x e Hib4.

Cepas	ng/uL	260/280	260/230	Volume total (μL)	Total (μg)
J1a12	86,24	1,91	2,25	200	17,24
9a5c	81,71	1,80	2,05	100	8,17
Pr8x	60,43	1,92	2,20	100	6,04
Hib4	88,92	2,02	2,29	100	8,89

O DNA plasmidial purificado das quatro cepas foi submetido a eletroforese longa em baixa voltagem (Saillard *et al.*, 2008), segundo protocolo descrito no item 3.6.2. O perfil da eletroforese mostrados na Figura 14A sugere a existência de pelo menos dois plasmídeos na cepa J1a12 pXF51 e pXF27. Por outro lado, Hib4 e Pr8x parecem possuir pelo menos um plasmídeo que apresentam diferentes mobilidades no gel. A partir da comparação da migração das bandas no gel sugerimos que a cepa Pr8x contenha um plasmídeo de tamanho similar ao pXF51, enquanto a cepa Hib4, aparentemente parece não apresentar nem o pXF51 nem o pXF27, e sim provavelmente outro plasmídeo distinto.

A figura 14B mostra uma eletroforese convencional em gel de agarose 0,8% com as duas novas amostras de extrações de DNA plasmidial das cepas 9a5c e J1a12, mostrando o padrão de bandas e confirmando a co-existência de pelo menos dois tipos de plasmídeos em J1a12, como predito pelo pirosequenciamento.

Figura 14. Análise eletroforética do DNA plasmidial das cepas J1a12, 9a5c, Pr8x e Hib4. **A.** Amostras de DNA plasmidial total isolados das cepas J1a12, Pr8x, Hib4 e 9a5c, respectivamente canaletas 2, 3, 4 e 5, foram submetidas a uma eletroforese longa em baixa voltagem em gel de agarose 0,5%. O marcador *Lambda DNA HindIII Digest* (New England Biolabs) está na canaleta 1. **B.** Amostras de DNA plasmidial total das cepas J1a12 e 9a5c, respectivamente canaletas 2 e 3, foram submetidas a uma eletroforese convencional em gel de agarose 0,8%. Os marcadores de tamanho 1kb (Fermentas) e *Lambda DNA HindIII Digest* (New England Biolabs) foram aplicados, respectivamente, nas canaletas 1 e 4.

A presença de pXF27 circular na cepa J1a12 foi confirmada em ensaios de PCR com os par de *primers* pXF27_circ planejado a partir da sequência predita para pXF27 de modo a originar um produto de ~3000 pb (Figura 15, canaleta 3). Como

esperado, pXF27 parece ser exclusivo da cepa J1a12, como mostrado pela ausência deste produto em ensaios de PCR realizados com DNA total isolado das cepas 9a5c, Hib4 e Pr8x (Figura 15). Este resultado foi confirmado com um segundo par de primers (pXF27_8) para região da CDS VirD4 e que gera um *amplicon* de 1181pb apenas na PCR do DNA isolado de J1a12 (Figura 15, canaleta 9).

Figura 15. Análise da presença de pXF27 nas cepas J1a12, 9a5c, Hib4 e Pr8x.

Os produtos resultantes de ensaios de PCR com os pares de *primers* pXF27_circ (canaletas 2 a 6) e pXF27_8 (8 a 12) a partir de DNA total extraído das cepas J1a12 (canaletas 3 e 9), Pr8x (canaletas 4 e 10), Hib4 (canaletas 5 e 11) e 9a5c (canaletas 6 e 12) foram submetidos a eletroforese em gel de agarose 1%. O marcador de tamanho 1kbp (Fermentas) foi aplicado nas canaletas 1 e 7. Os controles negativos das PCRs (sem adição de DNA molde) foram aplicados nas canaletas 2 e 8.

A presença de pXF51 circular nas cepas J1a12 e 9a5c também foi verificada com a utilização dos pares de *primers* pXF51_circ em ensaios de PCR utilizando-se DNA total purificado (Figura 16). Esta figura também mostra a confirmação adicional da presença de pXF51 obtida em ensaios de PCR com o par de *primers* pXF51_vapD. Ambos os pares de *primers* foram então utilizados para avaliar a presença de plasmídeos similares a pXF51 nas cepas Hib4 e Pr8x. Os resultados apresentados na figura 16 indicam que Hib4 parece não conter pXF51 embora possua um plasmídeo que contém sequência reconhecida pelo segundo par de *primers*. Por outro lado o resultado foi positivo para Pr8x, indicando a existência de um plasmídeo similar a pXF51 nesta cepa. Como será discutido adiante, o sequenciamento genômico destas cepas confirmou estas observações.

Figura 16: Análise da presença de pXF51 nas cepas J1a12, 9a5c, Hib4 e Pr8x. Os produtos resultantes de ensaios de PCR com os pares de *primers* pXF51_circ (canaletas 2 a 6) e pXF51_vapD (8 a 12) a partir de DNA total extraído das cepas J1a12 (canaletas 3 e 9), Pr8x (canaletas 4 e 10), Hib4 (canaletas 5 e 11) e 9a5c (canaletas 6 e 12) foram submetidos a eletroforese em gel de agarose 1%. O marcador de tamanho 1kbp (Fermentas) foi aplicado nas canaletas 1 e 7. Os controles negativos das PCRs (sem adição de DNA molde) foram aplicados nas canaletas 2 e 8.

4.4.4. Análise de pXF27 e pXF51 de J1a12 com enzimas de restrição

Com o objetivo de confirmarmos as montagens *in silico* das sequências de ambos os plasmídeos, pXF51 e pXF27, identificados no sequenciamento do genoma da cepa J1a12, realizamos diversas digestões com diferentes enzimas de restrição tanto do plasmídeo purificado como de produtos de amplificação de sequência selecionadas do plasmídeo.

A Figura 17A mostra a digestão do produto de PCR com *primers* pXF27_circ (fragmento de 3000 pb) com a enzima de restrição *SacII*. Conforme esperado, a digestão do fragmento de 3000 pb com esta enzima resultou em dois fragmentos com os tamanhos previstos pelo o mapa de restrição (uma banda de 1926 pb e outra de 1074 pb), indicando que a montagem *in silico* para este segmento parece estar correta.

Figura 17. Digestões de pXF27 com enzimas de restrição. **A.** Produto de PCR (3000pb) obtido de pXF27 foi digerido (canaleta 2 ou não (canaleta 3) com *SacII*. **B.** DNA plasmidial isolado de J1a12 (canaleta 2) e 9a5c (canaleta 3) foi digerido com *XhoI*. O DNA plasmidial não digerido de J1a12 e 9a5c foi aplicado, respectivamente, nas canaletas 4 e 5. Os marcadores de tamanho *DNA HindIII Digest* (New England Biolabs) e 1kb (Fermentas) foram aplicados nas canaletas 1 e 6. As imagens mostradas em A e B, correspondem a eletroforese em gel de agarose 1%. **C.** DNA plasmidial isolado de J1a12 (canaleta 2) e 9a5c (canaleta 3) foi digerido com *NdeI* e também com *SacII*, J1a12 (canaleta 4) e 9a5c (canaleta 5). O DNA plasmidial não digerido de J1a12 e 9a5c foi aplicado, respectivamente, nas canaletas 6 e 7. Os marcadores de tamanho *DNA HindIII Digest* (New England Biolabs) e 1kb (Fermentas) foram aplicados nas canaletas 1 e 8. As amostras foram submetidas a eletroforese em gel de agarose longa em baixa voltagem. A posição de migração de pXF27, de pXF51 de alguns marcadores de tamanho está indicada.

A partir da sequência proposta para pXF27 e pXF51 obtivemos os seus respectivos mapas de restrição (Davis, 2009) para seleção das enzimas de restrição a serem utilizadas. As enzimas utilizadas para os ensaios de digestão foram *XhoI*, *NdeI* e *SacII*.

A enzima de restrição *XhoI* apresenta quatro sítios de reconhecimento no pXF51 e nenhum no pXF27. Assim o esperado seria a detecção de quatro bandas (20701 pb, 17471 pb, 12071 pb e 915 pb) derivadas de pXF51. Observamos na Figura 17B as três bandas para o pXF51 na amostra de DNA plasmidial da cepa 9a5c (exceto pela ausência da menor banda, a qual provavelmente deve ter saído do gel nesta eletroforese) e quatro bandas na amostra de DNA plasmidial da cepa J1a12, três correspondendo ao pXF51 e uma ao pXF27 intacto.

A enzima de restrição *NdeI* apresenta três sítios de reconhecimento no pXF51 e dois no pXF27 e o esperado seria a detecção de três bandas (21180 pb, 18533 pb e 11445 pb) correspondentes ao primeiro e duas bandas correspondentes ao

segundo (26139 pb e 1119 pb). Observamos três bandas para o pXF51 na amostra de DNA plasmidial da cepa 9a5c e quatro bandas na amostra de DNA plasmidial da cepa J1a12, três correspondendo ao pXF51 e uma ao pXF27 (novamente, a menor banda provavelmente tenha saído do gel nesta eletroforese longa) (Figura 17C).

A enzima de restrição *SacII* apresenta dois sítios de reconhecimento no pXF51 e um no pXF27 e o esperado seria detectarmos duas bandas (32329 pb e 18829 pb) derivadas de pXF51 e uma banda correspondente a pXF27 linearizado. Observamos duas bandas para o pXF51 na amostra de DNA plasmidial da cepa 9a5c e três bandas na amostra de DNA plasmidial na cepa J1a12, duas correspondendo ao pXF51 e o pXF27 linearizado (Figura 17C).

Em resumo, todos os ensaios de digestão realizados indicam que as montagens *in silico* para ambos os plasmídeos estão corretas.

4.4.5. Análise da presença pXF27 em outros isolados de *X. fastidiosa*

Com o objetivo de tentar traçar um perfil da distribuição deste novo plasmídeo, denominado pXF27, em isolados de *X. fastidiosa* de diversos hospedeiros de diversas localizações geográficas, avaliamos sua provável presença através de ensaios de PCR com os *primers* pXF27_circ, que gera produtos de 3000 pb. Analisamos 235 amostras de DNA extraído de diversos isolados. A figura 18 exemplifica o resultado dos ensaios de PCR seguido de eletroforese convencional para alguns dos isolados. A Tabela 11 resume os resultados obtidos, indicando o número de isolados que foram analisados de cada localidade e quantos foram positivos ou negativos quanto à detecção do produto de 3000 pb.

Os isolados de *X. fastidiosa* foram obtidos de diferentes cidades, em quatro estados brasileiros (Minas Gerais, Paraná, São Paulo e Sergipe), além de um isolado da Argentina (município de Bella Vista, província de Corrientes). Aproximadamente 45% dos isolados foram positivos quanto à presença do pXF27 e foram encontrados em 67% das cidades, evidenciando que trata-se de um plasmídeo que está presente nos pomares do país, provavelmente sendo disseminado por conjugação entre as diferentes cepas que podem co-existir dentro do mesmo hospedeiro vegetal ou do inseto vetor.

Em um estudo anterior utilizando microarranjos de DNA e hibridização subtrativa, o genoma de seis linhagens de *X. fastidiosa* foi comparado (quatro linhagens isoladas de árvores com CVC no Estado de São Paulo, uma de uma árvore não-sintomática da mesma região e uma linhagem da Argentina que causa doença em laranjeira) possibilitando a identificação de sequência exclusivas destas

linhagens, sendo que um terço de elementos originários de outros microorganismos era homólogo a ORFs identificadas em *Verminephrobacter eiseniae* (da Silva *et al.*, 2007). Entre as linhagens analisadas por estes autores estão as cepas 9.12c (isolada de Gavião Peixoto), a cepa 36f (isolada de Ubarana) e a cepa Fb7 (isolada de Bella Vista, Corrientes, Argentina) que nós também avaliamos quanto à presença de pXF27. Interessantemente, a cepa 9.12c está entre aquelas que apresentaram um resultado positivo para a detecção do pXF27 que, como já mencionado, contem sequências similares ao plasmídeo pVEIS01 *V. eiseniae*, corroborando o estudo acima mencionado.

Figura 18. Detecção de pXF27 com par de primers pXF27_circ em diferentes isolados de *X. fastidiosa*. Os produtos resultantes de ensaios de PCR com o par de primers pXF27_circ (canaletas 3 a 20) a partir de DNA total extraído, respectivamente da cepa J1a12 e isolados P16a, P16b, P16c, P16d P16e, P16f, 2605a, 2605b, 2605c, 2605d, 2605e, G.P.7.4a, G.P.7.4b, G.P.7.4c, G.P.7.4d G.P.7.4e e G.P.7.4f foram submetidos a eletroforese em gel de agarose 1%. O marcador de tamanho 1kbp (Fermentas) foi aplicado na canaleta 1. O controle negativo da PCR (sem adição de DNA molde) foi aplicado na canaleta 2. A seta aponta o produto de ~3000 pb.

Tabela 11. Detecção de pXF27 em isolados de *X. fastidiosa* de diferentes hospedeiros obtidos em diferentes localidades.

	Cidade	+	-	total	Porcentagem
1	Paulínia	6	0	6	100
2	Ubarana	42	91	133	31,6
3	Gavião Peixoto	30	8	38	78,9
4	Neves Paulista	2	10	12	16,7
5	Santa Rita	17	8	25	68
6	Bebedouro	1	0	1	100
7	Itapetininga	1	0	1	100
8	Loanda (PR)	0	1	1	0
9	São Pedro do Paraná	0	1	1	0
10	Sergipe	0	1	1	0
11	Sete Lagoas (MG)	3	1	4	75
12	Cordeirópolis	0	1	1	0
13	Matão	1	0	1	100
14	Campinas	0	3	3	0
15	Corrientes (Argentina)	0	1	1	0
16	Cepa videira	0	1	1	0
17	Cepa ameixeira	0	1	1	0
18	ND	2	2	4	50

As colunas (+) e (-) indicam o número de isolados que foram, respectivamente, positivos ou negativos para a presença do pXF27. ND indica origem desconhecida. Os isolados das linhas 1 a 15 e 18 são, em sua maioria, de citros, porém incluem cepas de café e *Hibiscus*. Na análise foram incluídos um isolado de ameixeira e outro de videira. Todas as amostras de DNA já purificado foram cedidas pelo Dr. Helvécio Della Colleta Filho (Centro de Citricultura Sylvio Moreira, Cordeirópolis, SP).

4.5. Sequenciamento do genoma das cepas Pr8x e Hib4

Nós realizamos o pirosequenciamento do genoma das cepas Pr8x (de ameixeira) e Hib4 (de hibisco) no CATG/IQUSP utilizando a estratégia de *shotgun* (Margulies *et al.*, 2005; Ahmadian *et al.*, 2006) e um novo protocolo (*Rapid Library Preparation Method*) para preparação das bibliotecas. Como já mencionado, as preparações de DNA para pirosequenciamento devem apresentar razões de A260nm/230nm e A260nm/280nm entre 1,8 e 2,0 além de integridade adequada. A figura 19 mostra a integridade do DNA total purificado destas cepas, avaliada através de eletroforese convencional em gel de agarose e por eletroforese capilar no equipamento *Agilent 2100 Bioanalyzer* (Agilent Technologies) antes de serem sequenciadas.

Figura 19. Avaliação da integridade do DNA total purificado das cepas Pr8x e Hib4. A imagem à esquerda mostra a eletroforese convencional em gel de agarose do DNA purificado das cepas Pr8x e Hib4, respectivamente nas canaletas 2 e 4. O marcador de tamanho de 1kb (Fermentas) foi aplicado nas canaletas 1 e 3. Os perfis mostrados à direita correspondem à análise do DNA purificado das cepas Pr8x e Hib4 no Bioanalyzer. As bandas verde e lilás correspondem aos marcadores de tamanho, respectivamente 50 pb e 10380 pb.

O DNA total purificado foi utilizado para preparação de bibliotecas de ssDNA segundo o protocolo *Rapid Library Preparation* descrito no item 3.9.3.2 variando-se entretanto o tempo de nebulização (1 minuto e 10 segundos para o DNA de Pr8x e 2 minutos para o DNA da cepa Hib4) de modo a obter fragmentos na faixa de tamanho desejada (700-800pb) como mostrado na Figura 20.

Figura 20. Análise em *Bioanalyzer* das bibliotecas de ssDNA após nebulização. **A.** Análise da amostra de DNA de Pr8x após nebulização por 1 minuto e 10 segundos. **B.** Análise da amostra de DNA de Hib4 após nebulização por 2 minutos. As bandas verde e lilás correspondem, respectivamente, aos marcadores de tamanho de 35 pb e 10380 pb.

Como descrito no item 3.8.2, os demais passos do protocolo incluem a quantificação da biblioteca de ssDNA por fluorescência, titulação das bibliotecas por PCR em emulsão (emPCR) e realização da emPCR em maior escala (emPCR definitiva). Após esta etapa é realizada a etapa de enriquecimento na qual são removidos *beads* que tiveram seus processos de amplificação mal sucedidos. As *beads* são novamente contadas e as porcentagens de enriquecimento são calculadas as quais, segundo o protocolo, devem estar na faixa de 5-20%. Entretanto, ambas as bibliotecas tiveram suas porcentagens de enriquecimento após emPCR definitiva acima dos valores recomendados (para a cepa Pr8x foi de 45% e para a cepa Hib4 de 28%), indicativo da existência de *beads* mistas, ou seja, que contenham mais de um fragmento de ssDNA ligados a uma única *bead* antes

do processo de emPCR. Ainda assim, optamos prosseguir com o processo de pirossequenciamento, realizando duas corridas no equipamento 454 em placas PTP (*PicoTiterPlate*) divididas em duas regiões, sendo uma região destinada ao sequenciamento de Hib4 e outra para Pr8x. Os poços contendo as *beads* mistas produzem sinais luminosos distintos simultaneamente e serão excluídos na etapa de processamento das imagens, reduzindo o número de *reads* gerados por corrida.

A tabela 12 resume os dados obtidos nas duas corridas realizadas. Embora, o total de *reads* de cada corrida individual tenha ficado até 60% abaixo do esperado, os tamanhos médios dos *reads* gerados (~430 pb) apresentaram valores considerados excelentes e dentro do esperado pelo protocolo que utilizamos. Os gráficos apresentados na Figura 21 mostram a distribuição de *reads* obtidos com o pirossequenciamento das cepas Pr8x e Hib4, com picos entre 400-500 pb e uma baixa quantidade de sequências menores. Vale ressaltar que mudanças nos reagentes e protocolo recentemente introduzidas pela 454/Roche resultam na obtenção de sequências com tamanho médio de 700-800 pb (Dra. Layla Farage Martins, CATG-IQUSP, comunicação pessoal).

Tabela 12. Resumo do desempenho do pirossequenciamento das cepas Pr8x e Hib4

Cepa	Total de reads	Tamanho médio dos reads (bases)	Tamanho do menor/maior read (bases)	Total de bases	Cobertura estimada do genoma
Pr8x ¹	199.000	415	40,48/1194,1	82 Mpb	30 vezes
Hib4 ¹	298.000	440	40,61/1193,1	131 Mpb	48,5 vezes
Pr8x ²	208.000	426	40,46/840,1	89 Mpb	33 vezes
Hib4 ²	315.000	435	40,73/818,1	137 Mpb	51 vezes

O total de *beads* obtido após a etapa de enriquecimento foi suficiente para realização de duas corridas. O desempenho da primeira corrida¹ está apresentado nas duas primeiras linhas e da segunda² nas duas últimas linhas da tabela. A cobertura foi estimada com base no tamanho do genoma da cepa 9a5c (~2,7 Mpb).

Figura 21. Gráficos de distribuição dos tamanhos dos *reads* obtidos no pirossequenciamento. Foram realizadas duas corridas, sendo que em cada uma delas metade da placa foi utilizada para a biblioteca de Pr8x (**A**) e Hib4 (**B**). Os gráficos são fornecidos pelo diretamente pelo *software* do equipamento de sequenciamento 454/Roche, após o processamento das imagens e conversão em sequências.

4.5.1. Montagem dos genomas de Pr8x e Hib4

As montagens preliminares dos genomas das cepas Pr8x e Hib4 foram realizadas utilizando o programa Newbler que acompanha a plataforma 454/Roche e o total de *reads* gerados nas duas corridas realizadas. A tabela 13 resume os resultados das montagens do genoma de Pr8x (92 *contigs*) e Hib4 (115 *contigs*), os quais foram superiores ao obtido na montagem preliminar do genoma da cepa J1a12 (374 *contigs*, item 4.3.1), refletindo a melhor qualidade das sequências obtidas no sequenciamento de Pr8x e Hib4.

Na tentativa de obter um único *contig* na montagem destes genomas, foram utilizados os programas MIRA (http://chevreux.org/projects_mira.html) e Cross_Match (<http://www.phrap.org/phredphrap/general.html>) juntamente com intervenções manuais. Estas montagens, ainda preliminares (versão 1.0), resultaram em um cromossomo principal de 2.666.242 pb para cepa Pr8x e um cromossomo principal de 2.813.297 pb para a cepa Hib4, além de um plasmídeo de

~39 kpb para a primeira cepa e um de ~64 kpb para a segunda. Os resultados destas montagens foram submetidos para anotação automática na plataforma IMG/ER e os dados obtidos estão sumarizados na Tabela 14. Se esta montagem for confirmada, a cepa Hib4 apresenta, até o momento, o maior cromossomo principal já sequenciado de cepas de *X. fastidiosa*, assim como o maior plasmídeo já encontrado nesta espécie. Assim como no caso da cepa J1a12, os genomas de Pr8x e Hib4 apresentam características gerais semelhantes às dos genomas de outras cepas de *X. fastidiosa* já sequenciadas.

Tabela 13. Resultados das montagens dos genomas das cepas Pr8x e Hib4 com o programa *Newbler*

Cepa	Número de <i>reads</i> na montagem	Número de bases na montagem	Total de <i>contigs</i> gerados	Tamanho do maior <i>contig</i> (bases)	Tamanho médio dos maiores <i>contigs</i> (bases)	Total de bases na montagem
Pr8x	402.486	~170x10 ⁶	92	260.475	42.370	2.635.224
Hib4	605.316	~266x10 ⁶	115	191.698	30.944	2.731.097

Tabela 14. Características gerais dos genomas das cepas Pr8x e Hib4

Características	Pr8x		Hib4	
	Cromossomo	pXF39	Cromossomo	pXF64
Tamanho	2.666.242 pb	39.580 pb	2.813.297	64.251
G+C (%)	52.63%	49.55%	52,70%	52,3%
Genes codificadores de proteínas	2598	57	2761	67
Genes codificadores de proteínas com função predita	1733 (65.15%)	29 (50.88%)	1839 (65,17%)	19 (27,36%)
Genes de rRNA	6	-	6	-
rRNA 5S	2	-	2	-
rRNA 16S	2	-	2	-
rRNA 23S	2	-	2	-
tRNA genes	51	-	51	-
Outros genes de RNA	7	-	6	-
tmRNA	1	-	1	-
SNP-RNA	1	-	1	-
RNase P-RNA	1	-	1	-
intron Grupo II -RNA	1	-	-	-
6S/SsrS RNA	1	-	1	-
yybP-ykoY líder	1	-	1	-
TPP <i>riboswitch</i>	1	-	1	-
CDS com ortólogos em 9a5c				
≥80% identidade/ Valor E ≤10 ⁻¹⁰	2.157 (83,0%)		2.225 (80,5%)	
≥60% identidade/ Valor E ≤10 ⁻¹⁰	2.205 (84,8%)		2.285 (82,7%)	
≥30% identidade/ Valor E ≤10 ⁻¹⁰	2.216 (85,3%)		2.312 (83,7%)	
CDS com ortólogos em Temecula				
≥80% identidade/ Valor E ≤10 ⁻¹⁰	1.890 (72,7%)		1.977 (71,6%)	
≥60% identidade/ Valor E ≤10 ⁻¹⁰	1.967 (75,7%)		2.033 (73,6%)	
≥30% identidade/ Valor E ≤10 ⁻¹⁰	2.002 (77,0%)		2.083 (75,4%)	

A comparação entre as sequências de nucleotídeos do cromossomo principal de Pr8x e Hib4 com o cromossomo de 9a5c indica a ocorrência de inversões e rearranjos de várias regiões genômicas (Figura 22) assim como observado na análise comparativa com o genoma de J1a12. Interessantemente, os genomas de Pr8x e Hib4 aparentemente são mais colineares entre si do que quando comparados a 9a5c ou J1a12. Estas análises serão refinadas após a obtenção de novas versões de montagens destes genomas.

Figura 22. Alinhamento do cromossomo principal das cepas Pr8x, Hib4, J1a12 e 9a5c. Os diagramas do tipo “Dotplot” foram gerados a partir da comparação da sequência de nucleotídeos do cromossomo principal das cepas Pr8x, Hib4, J1a12 e 9a5c com o programa MUMmer (Delcher *et al.*, 2003) disponível na plataforma IMG/ER, conforme indicado. As regiões de similaridade encontradas na mesma orientação estão indicadas em azul e em vermelho estão indicadas regiões de similaridade em orientação oposta (inversão).

4.5.2. Análises comparativas dos genomas de Pr8x e Hib4

Para as análises comparativas das primeiras montagens dos genomas de Pr8x e Hib4 utilizamos a anotação automática feita tanto no servidor RAST (Aziz et al., 2008) como na plataforma IMG-ER (Markowitz *et al.*, 2012) com estratégia semelhante à aplicada nas análises comparativas do genoma da J1a12 e descrita anteriormente (seção 4.3.2).

4.5.2.1. Análises comparativas do genoma de Pr8x

A cepa Pr8x isolada de ameixeira apresenta um fenótipo semelhante à cepa 9a5c em cultura *in vitro*, e desta forma, suspeitávamos que ela tivesse uma maior similaridade em relação ao seu conteúdo gênico em comparação com a cepa de citros 9a5c. Corroborando esta hipótese, apenas algumas diferenças importantes foram detectadas e estão resumidas na Tabela 15.

Entre as CDS ausentes em Pr8x, reflexo de deleções em sua sequência genômica, estão XF1250 (arginina deaminase) e XF1668 (proteína hicB) também ausentes na cepa J1a12. Especulamos se também em Pr8x, a CDS XFPR8X_00006380 (amido-hidrolase) poderia compensar a ausência da arginina deaminase que também hidrolisa ligações carbono-nitrogênio não peptídicas.

Genes do sistema toxina/antitoxina higB/higA (Hurley e Woychik, 2009; Fivian-Hughes e Davis, 2010), respectivamente XF0721 e XF0720, presentes em uma região correspondente a um fago em 9a5c, também estão ausentes em Pr8x.

Além disso, um *cluster* de CDS (XF2045 a XF2060) relacionadas a proteínas de transferência por conjugação (genes *trb*) também está ausente. Em 9a5c, este *cluster* é aparentemente originário de pXF51, ausente em Pr8x. Por outro lado, esta cepa possui um plasmídeo pXF39 que inclui ortólogos destas CDS.

Entre as CDSs com alterações que possivelmente afetam a função do produto codificado, estão CDSs com funções de adesão tais como uma cópia da adesina PilY1 do pilus do tipo IV (XF0032), as três cópias para proteínas “*hemagglutinin-like*” (XF0889, XF2196 e XF2775) além de uma proteína de superfície, a XF1529, descrita como uma adesina não-fimbrial (Caserta *et al.*, 2010). Assim como, no caso discutido para J1a12, não podemos excluir a ocorrência de erros de montagem na sequência das CDS que codificam proteínas similares a hemaglutininas. As demais alterações, em particular na XF1529, parecem não prejudicar a adesão célula-célula (agregados similares aos de 9a5c são observados na cultura *in vitro* de Pr8x) nem a adesão ao substrato (deposição de células no biofilme na interface ar-líquido na superfície de vidro do frasco de cultura), como mostrado na Figura 1. Outra CDS

com possível alteração de função seria a XF0373 que codifica para uma proteína que atua na montagem do pilus do tipo IV. Ela apresenta uma deleção de um nucleotídeo que gera um códon de parada prematuro, gerando uma proteína truncada. Este fato corroboraria com a observação das colônias de Pr8x que não exibem uma “franja” correspondente às células satélites se movimentando por *twitching motility*, quando em cultivo em meio sólido PWG (Tabela 3). Assim como em outras bactérias, em *X. fastidiosa*, a motilidade *twitching* depende do pilus do tipo IV (Mattick, 2002; Meng *et al.*, 2005). Entretanto, esta deleção está localizada em uma região homopolimérica, podendo ser resultado de erros no pirossequenciamento. Portanto, novas análises necessitam ser realizadas para a confirmação destes resultados.

Observamos a ocorrência de deleções nas regiões de CDS que codificam hemolisinas (XF0668, XF1011, XF2407 e XF2759). A função destas proteínas na virulência de *X. fastidiosa* ainda permanece incógnita e portanto não é possível fazer alguma associação destas alterações com fenótipo de Pr8x. De qualquer forma, esta cepa foi isolada de ameixeira com sintomas de escaldadura da folha, mas nenhuma outra confirmação da virulência desta linhagem foi realizada até o momento.

Finalmente, é importante destacar que o genoma de Pr8x codifica a proteína precursora de poligalacturonase completa (ortóloga de XF2466 em 9a5c e PD1485 em Temecula). Na cepa 9a5c e em outras cepas isoladas de citros e de cafeeiros, este gene apresenta uma inserção de um nucleotídeo que resulta no aparecimento de um códon de terminação prematuro, sendo que a ausência desta enzima tem sido associada a um fenótipo menos agressivo para estas cepas (Van Sluys *et al.*, 2003; Chatterjee *et al.*, 2008). Assim, essa enzima, que participa da degradação da parede vegetal das fissuras membranares entre vasos do xilema e consequente espalhamento da bactéria em videiras (Roper *et al.*, 2007; Pérez-Donoso *et al.*, 2010), é aparentemente funcional na cepa Pr8x, como mostra o alinhamento apresentado na Figura S1 (no apêndice). Interessantemente, estudo anterior com uma linhagem de ameixeira (*plum 1*), foi observado que o seu perfil genômico geral sugere que tenha derivado diretamente de isolados da América do Norte (Nunes *et al.*, 2003), os quais codificam a poligalacturonase íntegra. Como será discutido adiante, também a cepa Hib4 (isolada de hibisco) apresenta o gene de poligalacturonase intacto.

Tabela 15. CDS de 9a5c divergentes no cromossomo de Pr8x

CDS em 9a5c	Produto predito	Característica em Pr8x
XF0032	<i>PilY1 (Tfp pilus assembly protein, tip-associated adhesin PilY1)</i>	Com deleções, inserções e polimorfismos/ códon de parada prematuro
XF0373	<i>fimbrial assembly protein (pilQ)</i>	Deleção de 1 nt* /códon de parada prematuro
XF0668	<i>hemolysin-type calcium binding protein</i>	Inserções, deleções e polimorfismos/ CDS com N-terminal menor
XF0720	<i>proteína HigA (antitoxina de HigB)</i>	Ausente
XF0721	<i>proteína toxina HigB</i>	Ausente
XF0889	<i>hemagglutinin-like secreted protein (pspA)</i>	Deleções/CDS íntegra ausente
XF1011	<i>hemolysin-type calcium binding protein</i>	Deleções/CDS menor tamanho
XF1026	<i>serine protease (pspB)</i>	Inserções e polimorfismos/ CDS com N-terminal menor
XF1250	<i>arginine deaminase (rocF)</i>	Ausente
XF1529	<i>surface protein (hsf)</i>	Deleção/CDS sem região central de 669 aminoácidos
XF1668	<i>HicB-related protein</i>	Ausente
XF1851	<i>serine protease</i>	Deleções/ altera N-terminal/elimina peptídeo sinal
XF1948	<i>colicin V production protein</i>	Inserção*/altera C-terminal
XF2045 a XF2060	<i>conjugal transfer proteins</i>	Ausentes
XF2196	<i>hemagglutinin-like secreted protein (pspA)</i>	Deleções/CDS íntegra ausente
XF2297	<i>DNA methylase</i>	Inserção/CDS maior
XF2341	<i>heat shock protein GrpE</i>	CDS com N-terminal menor
XF2407	<i>bacteriocin/ hemolysin</i>	Polimorfismos e deleções reduzem tamanho da proteína predita de 2407 para 1216 aminoácidos
XF2466	<i>polygalacturonase precursor</i>	CDS completa sem frameshift
XF2723	<i>type I restriction-modification system DNA methylase</i>	Estas duas sequências estão fusionadas em J1a12 (XFPR8X_00025700) assim como nas cepas Temecula (PD2072) e J1a12
XF2724	<i>type I restriction-modification system</i>	
XF2759	<i>hemolysin-type calcium binding protein</i>	Inserções e polimorfismos/CDS com C-terminal mais curto
XF2775	<i>hemagglutinin-like secreted protein</i>	Deleções/CDS íntegra ausente

*O asterisco indica que inserção ou deleção de ocorreu em sequência homopolimérica.

4.5.2.2. Análises comparativas de pXF39

O sequenciamento do genoma da cepa Pr8x revelou a existência de um provável plasmídeo com tamanho de 39.580 pb, que denominamos pXF39. Este plasmídeo tem similaridades com pXF51 encontrado nas cepas 9a5c (Marques *et al.*, 2001) e J1a12 (item 4.4.1) ainda que seja um pouco menor (Figura S2 no apêndice). Em estudo anterior com a linhagem *plum I*, também isolada de ameixeira, foi observado, utilizando-se hibridização de microarranjos de DNA, que esta cepa continha sequências de quase todas as ORFs do pXF51 (Nunes *et al.*, 2003; da Silva *et al.*, 2007). A tabela 16 lista as CDS de pXF39 e seus respectivos ortólogos em pXF51 da cepa 9a5c. Várias CDS hipotéticas do plasmídeo pXF51 estão ausentes em pXF39 refletindo a deleções e rearranjos de um segmento de ~12000pb e que inclui CDS de genes *trb* (XFa0036, XFa0037, XFa0039, XFa0040, XFa0041, XFa0042, XFa0043 e XFa0044) relacionados à conjugação. Contudo, todo o sistema VirB/VirD4, de formação do sistema de secreção do tipo IV (Christie *et al.*, 2005) está presente e sem deleções em pXF39, sugerindo que este é um plasmídeo conjugativo.

Além disso, duas proteínas de estabilização e manutenção do pXF51 também parecem estar ausentes em pXF39, XFa0029 (*parE*) e XFa0027 (toxina *pemK*, do sistema toxina-antitoxina PemK/PemI). Entretanto, o sistema toxina/antitoxina YdcE/YdcD, codificados pelas CDSs XFa0022 e XFa0023, respectivamente, está completo.

Além das CDS ortólogas de pXF51, também verificamos que pXF39 contém um conjunto de CDS que não está presente em pXF51. Dentre elas, se encontra mais um sistema de toxina-antitoxina, YafQ/DinJ, cuja sequência de nucleotídeos está presente no plasmídeo pXF5823, obtido da linhagem LAR20 isolada de laranjeira doce com sintomas de CVC e, utilizado como vetor para transformação de *X. fastidiosa* (Qin e Hartung, 2001).

Tabela 16. CDS do plasmídeo pXF39 da cepa Pr8x e seus produtos preditos

CDS pXF39	Produto predito	Ortólogo em pXF51 da cepa 9a5c
1	<i>transcriptional regulator</i>	XFa0001
2	<i>virB1</i>	XFa0002
3	<i>traB</i>	XFa0003
4	<i>hypothetical protein</i>	XFa0004
5	<i>virB2</i>	XFa0005
6	<i>virB3</i>	XFa0006
7	<i>virB4</i>	XFa0007
8	<i>virB5</i>	XFa0008
9	<i>hypothetical protein</i>	XFa0009
10	<i>hypothetical protein</i>	XFa0010
11	<i>virB6</i>	XFa0011
12	<i>virB7</i>	XFa0065
13	<i>virB8</i>	XFa0012
14	<i>virB9</i>	XFa0013
15	<i>virB10</i>	XFa0014
16	<i>virB11</i>	XFa0015
17	<i>virD4</i>	XFa0016
18	<i>DNA primase TraO</i>	XFa0017
19	<i>hypothetical protein</i>	-
20	<i>YafQ toxin protein</i>	-
21	<i>DNA-damaged- inducible protein J</i>	-
22	<i>hypothetical protein</i>	-
23	<i>DNA-invertase</i>	XFa0019
24	<i>hypothetical protein</i>	XFa0020
25	<i>programmed cell death toxin YdcE</i>	XFa0022
26	<i>programmed cell death antitoxin YdcD</i>	XFa0023
27	<i>protein of unknown function DUF891</i>	XFa0024
28	<i>putative transcriptional regulator</i>	-
29	<i>hypothetical protein</i>	-
30	<i>Acetyltransferase</i>	-
31	<i>hypothetical protein</i>	-
32	<i>hypothetical protein</i>	-
33	<i>DNA nicking endonuclease TraR</i>	XFa0047
35	<i>hypothetical protein</i>	XFa0049
36	<i>stability protein StdB</i>	XFa0050
37	<i>hypothetical protein</i>	XFa0051
38	<i>virulence-associated protein D</i>	XFa0052
39	<i>hypothetical protein</i>	-
40	<i>hypothetical protein</i>	-
41	<i>hypothetical protein</i>	XFa0054
42	<i>hypothetical protein</i>	XFa0055
43	<i>hypothetical protein</i>	XFa0057

44	<i>hypothetical protein</i>	XFa0058
45	<i>plasmid replication protein ParB</i>	XFa0059
46	<i>plasmid replication protein</i>	XFa0060
47	<i>single-strand DNA-binding protein</i>	XFa0061
48	<i>protein of unknow funtion DUF3560</i>	XFa0062
49	<i>phage-related protein</i>	-
50	<i>hypothetical protein</i>	-
51	<i>hypothetical protein</i>	-
52	<i>hypothetical protein</i>	-
53	<i>phage-related protein</i>	-
54	<i>hypothetical protein</i>	-
55	<i>hypothetical protein</i>	XFa0064
56	<i>hypothetical protein</i>	-
57	<i>hypothetical protein</i>	-

4.5.2.3. Análises comparativas do genoma de Hib4

A tabela 17 lista as principais diferenças que verificamos quando comparamos o conjunto selecionado de CDS da cepa 9a5c com as CDS anotadas no genoma da cepa Hib4. Assim como verificado para cepa Pr8x e J1a12, entre as CDS ausentes em Hib4, reflexo de deleções em sua sequência genômica, estão XF1250 (arginina deaminase) e XF1668 (proteína HicB). Assim como, nestas duas cepas, também em Hib4, uma CDS que codifica uma amido-hidrolase (XFHIB4_00006310) poderia compensar a ausência da arginina deaminase.

Além destas duas CDS, também na cepa Hib4 verificamos a deleção de XF0077, que codifica um precursor de adesina de fimbrias e também está ausente na cepa J1a12. Como mostrado na Figura 1, o fenótipo de agregação celular de Hib4 é intermediário ao das cepas 9a5c (grandes agregados celulares) e J1a12 (células dispersas no meio), porém com pouquíssima adesão à superfície de vidro no cultivo em meio líquido. Além disso, XF1529 (adesina não-fimbrial) (Caserta *et al.*, 2010) também apresenta uma extensa deleção na sua porção central como verificado no genoma da cepa Pr8x. Se por um lado esta alteração parece não afetar o fenótipo agregativo da cepa Pr8x, podemos especular que ausência de duas adesinas possa estar correlacionada com o fenótipo de menor adesão de Hib4.

De modo semelhante ao observado para J1a12 e Pr8x, as CDS as três cópias para proteínas “hemagglutinin-like” (XF0889, XF2196 e XF2775) estão incompletas, e, também neste caso, não podemos excluir a ocorrência de erros de montagem na sequência das CDS que codificam proteínas similares a hemagglutininas, como discutimos anteriormente.

A montagem do genoma de Hib4 que analisamos também revela a ocorrência de deleções nas regiões de CDS que codificam hemolisinas (XF0668, XF1011, XF2407 e XF2759). Embora a cepa Hib4 tenha sido isolada de hibisco com sintomas de escaldadura da folha, não há, até o momento, outra confirmação da virulência desta linhagem.

As comparações realizadas mostraram diversas alterações na sequência de nucleotídeos do genoma de Hib4 (INDELs e polimorfismos) que aparentemente não sugerem alterações funcionais óbvias e que, portanto, não foram discutidas neste trabalho. A mesma afirmação é válida para as análises comparativas dos genomas de J1a12 e Pr8x. Contudo, destacamos a alteração observada na CDS ortóloga de XF1952, que codifica uma proteína quinase relacionada a quimiotaxia. Na cepa Hib4 esta CDS apresenta uma inserção de 21 nucleotídeos que não provoca mudança da fase de leitura (Figura S3 no apêndice). Esta inserção mostrou-se útil, pois foi confeccionado um par de *primers* para essa região, adequado para genotipar e diferenciar a cepa Hib4 das demais cepas em cultivo em nosso laboratório (Figura 3B).

Finalmente, é importante destacar que, assim como verificamos em Pr8x, o genoma de Hib4 também codifica a proteína precursora de poligalacturonase completa (Figura S1 no apêndice). Como já mencionado, esta enzima é importante para colonização do hospedeiro (Roper *et al.*, 2007; Pérez-Donoso *et al.*, 2010), sendo que cepas isoladas de citros e de cafeeiros têm uma mutação de *frameshift* neste gene o que tem sido associado seu fenótipo menos agressivo (Van Sluys *et al.*, 2003; Chatterjee *et al.*, 2008).

Tabela 17. CDS de 9a5c divergentes no cromossomo de Hib4

CDS em 9a5c	Produto predito	Característica em Hib4
XF0077	<i>fimbrial adhesin precursor</i>	Ausente
XF0668	<i>hemolysin-type calcium binding protein</i>	Inserções e polimorfismos/ CDS com N-terminal maior
XF0889	<i>hemagglutinin-like secreted protein (pspA)</i>	Deleções/CDS íntegra ausente
XF1011	<i>hemolysin-type calcium binding protein</i>	Deleções/CDS menor tamanho
XF1026	<i>serine protease (pspB)</i>	Inserções e polimorfismos/CDS com N-terminal maior
XF1250	<i>arginine deaminase (rocF)</i>	Ausente
XF1529	<i>surface protein (hsf)</i>	Deleção/CDS sem região central de ~650 aminoácidos
XF1668	<i>HicB-related protein</i>	Ausente
XF1774	<i>DNA methyltransferase</i>	Deleções/códon de parada prematuro
XF1851	<i>serine protease</i>	Deleções/ altera N-terminal
XF1948	<i>colicin V production protein</i>	Inserção*/ altera C-terminal
XF1952	<i>chemotaxis-related protein kinase</i>	Inserção de 21 nt em fase
XF2046	<i>conjugal transfer protein (trbL)</i>	Deleções/CDS menor
XF2048	<i>conjugal transfer protein (trbJ)</i>	Região com deleções e inserções
XF2053	<i>conjugal transfer protein (trbE)</i>	
XF2056	<i>conjugal transfer protein (trbB)</i>	
XF2059	<i>conjugal transfer protein (traE)</i>	Deleções/codon de parada prematuro
XF2196	<i>hemagglutinin-like secreted protein (pspA)</i>	Deleções/CDS íntegra ausente
XF2297	<i>DNA methylase</i>	Ausente
XF2341	<i>heat shock protein GrpE</i>	CDS com N-terminal menor
XF2407	<i>bacteriocin/ hemolysin</i>	Polimorfismos e deleções reduzem tamanho da proteína predita de 2407 para 1380 aminoácidos
XF2466	<i>polygalacturonase precursor</i>	CDS completa sem frameshift
XF2721	<i>type I restriction-modification system endonuclease</i>	Com deleções/CDS menor
XF2723	<i>type I restriction-modification system DNA methylase</i>	Estas duas sequências estão fusionadas em J1a12 (XFHIB4_00027620) assim como na cepa Temecula (PD2072)
XF2724	<i>type I restriction-modification system</i>	
XF2726	<i>type I restriction-modification system specificity determinant</i>	Sequência de nucleotídeos muito divergente porém codifica CDS ortóloga
XF2759	<i>hemolysin-type calcium binding protein</i>	Deleções/CDS íntegra ausente
XF2775	<i>hemagglutinin-like secreted protein</i>	Deleções/CDS íntegra ausente

*O asterisco indica que a inserção ou deleção ocorreu em sequência homopolimérica.

4.5.2.4. Análises comparativas de pXF64

Assim como outras cepas de *X. fastidiosa* já estudadas, também a cepa Hib4 parece ter o seu próprio plasmídeo, neste caso com tamanho de 64.251 pb e que denominamos pXF64.

Através de comparações por BlastN no NCBI, observamos que pXF64 tem uma cobertura de 71% com identidade de até 86% com o plasmídeo pBVIE04 de *Burkholderia vietnamensis* cepa G4 (*GenBank number*: CP000620.1). Trata-se de um plasmídeo inédito em cepas de *X. fastidiosa*. pXF64 também tem uma cobertura de 10% com até 99% de identidade com pXF51 da cepa 9a5c.

A tabela 18 lista as 67 CDS preditas pela anotação automática de pXF64 no servidor RAST e IMG-ER. Embora anotadas como hipotéticas, este plasmídeo parece conter todo o repertório de genes responsáveis por sua conjugação que são similares com genes *tra* (CDSs 3, 22, 24, 25, 26, 27, 53, 54, 55, 57, 59, 60, 61) e genes *trb*. Aparentemente, apresenta somente um sistema de toxina/antitoxina, Doc/Phd (Garcia-Pino *et al.*, 2008), além de genes (genes *par*) relacionados a segregação eficiente de cópias de plasmídeos após replicação dos mesmos e posterior divisão celular.

Tabela 18. CDS preditas pela anotação automática de pXF64

CDS pXF64	Produto predito		
1	<i>hypothetical protein</i>	28	<i>MobA</i>
	<i>UDP-3-O-[3-hydroxymyristoyl] glucosamine N- acyltransferase</i>	29	<i>hypothetical protein</i>
2	<i>acyltransferase</i>	30	<i>hypothetical protein</i>
3	<i>hypothetical protein</i>	31	<i>mobilization protein MobB</i>
4	<i>hypothetical protein</i>	32	<i>hypothetical protein</i>
	<i>uncharacterized protein conserved in bactéria (zeta toxin)</i>		<i>CobQ/ CobB/ MinD/ ParA nucleotide binding domain</i>
5	<i>DNA methylase</i>	33	<i>hypothetical protein</i>
6	<i>hypothetical protein</i>		<i>UDP-3-O-[3-hydroxymyristoyl] glucosamine N- acyltransferase</i>
7	<i>hypothetical protein</i>	35	<i>death on curing protein, Doc toxin</i>
8	<i>hypothetical protein</i>	36	<i>prevent host death protein, Phd antitoxin</i>
9	<i>hypothetical protein</i>	37	<i>probable extracellular repeat protein</i>
10	<i>hypothetical protein</i>	38	<i>DNA-invertase</i>
11	<i>DNA topoisomerase III</i>	39	<i>hypothetical protein</i>
	<i>conjugative transfer protein/ transglycosylase SLT domain</i>	40	<i>hypothetical protein</i>
12	<i>TrbA-like protein</i>	41	<i>hypothetical protein</i>
13	<i>conjugal transfer protein trbB</i>	42	<i>hypothetical protein</i>
14	<i>DNA primase</i>	43	<i>hypothetical protein</i>
15	<i>hypothetical protein</i>		<i>predicted transcriptional regulator</i>
16	<i>predicted ATPase</i>	44	<i>plasmid replication/partition protein</i>
	<i>phosphatidylserine/phosphat idylglycerophosphate/ cardiolipin synthases and related enzymes-like protein precursor</i>	44	<i>chromosome (plasmid) partitioning protein</i>
18	<i>hypothetical protein</i>	45	<i>phage-related protein</i>
19	<i>hypothetical protein</i>	46	<i>uncharacterized conserved small protein</i>
20	<i>hypothetical protein</i>	47	<i>hypothetical protein</i>
	<i>IncI1 plasmid conjugative transfer DNA primase</i>	48	<i>chromosome (plasmid) partitioning protein</i>
21	<i>IncI1 plasmid conjugative transfer protein TraL</i>	49	<i>hypothetical protein</i>
22	<i>hypothetical protein</i>	50	<i>replication initiator protein A</i>
23	<i>IncI1 plasmid conjugative transfer protein TraK</i>	51	<i>hypothetical protein</i>
24	<i>IncI1 plasmid conjugative transfer protein TraJ, related to pilus</i>	52	<i>plasmid conjugative transfer protein</i>
	<i>biogenesis/retracton protein</i>	53	<i>protein of unknown function (DUF3625)</i>
25	<i>IncI1 plasmid conjugative transfer protein TraI</i>	54	<i>hypothetical protein</i>
26	<i>IncI1 plasmid conjugative transfer protein TraH</i>	55 a	<i>hypothetical protein</i>
27		66	
		67	<i>superfamily I DNA and RNA helicases</i>

4.6. Análises filogenéticas de cepas de *X. fastidiosa*

Desde o primeiro sequenciamento de um genoma de *X. fastidiosa* (Simpson *et al.*, 2000) até os mais recentes genomas sequenciados (Bhattacharyya *et al.*, 2002; Van Sluys *et al.*, 2003; Chen *et al.*, 2010; Schreiber *et al.*, 2010; Zhang *et al.*, 2011), várias análises filogenéticas tem sido feitas a fim de entender a origem, especiação e diversidade genética deste fitopatógeno, que atualmente é encontrado na América do Norte, América Central e América do Sul. Os marcadores moleculares mais usados para estas análises são rRNA16S, a região ITS (*Internal Transcribed Region*) do operon de rRNA 16S-23S (devido a sua alta variação de comprimento e sequência) e um conjunto de genes *housekeeping* (Mehta e Rosato, 2001; Rodrigues *et al.*, 2003; Schuenzel *et al.*, 2005; Montero-Astúa *et al.*, 2007; Almeida *et al.*, 2008; Pieretti *et al.*, 2009).

No entanto, esses marcadores, principalmente o rRNA16S, não fornecem uma resolução dos ramos do cladograma que permitam definir claramente os isolados das Américas (Mehta e Rosato, 2001; Randall *et al.*, 2009). Essa indefinição das *taxa*, principalmente da América do Sul, se deve a alta conservação das sequências rRNA 16S a exemplo de outras bactérias (Asker e Ohta, 2002; Ludwig, 2007; Shneyer, 2007; Rooney *et al.*, 2009; McTaggart *et al.*, 2010).

A Análise de Sequências Multilocus (MLSA) é método de reconstrução filogenética adotado para obtenção de uma classificação taxonômica mais robusta de cepas proximamente relacionadas e com alto nível de recombinação genética (Maiden *et al.*, 1998; Hanage *et al.*, 2005). O princípio deste método consiste na concatenação de várias sequências dos mesmos genes para um grupo de organismos (espécies ou linhagens), resultando em uma "supermatriz", que é utilizada para construção da filogenia (Maiden *et al.*, 1998). Análises de sequências concatenerizadas de múltiplos *loci* são frequentemente utilizadas para obter uma visão mais robusta e abrangente das relações filogenéticas e foram usadas para identificar a subespécie *sandyi* como uma nova subespécie de *X. fastidiosa* (Randall *et al.*, 2009).

Para a reconstrução filogenética de *X. fastidiosa*, incluindo as cepas descritas neste trabalho, utilizamos grupos de sequências de sete genes *housekeeping* cada, previamente utilizados para reconstruções filogenéticas por MLSA de cepas de *X. fastidiosa* (Nunney *et al.*, 2012). As sequências de ortólogos destes genes foram recuperadas de todos os genomas de *X. fastidiosa* publicamente disponíveis no NCBI (cepa sul-americana 9a5c e as cepas norte-americanas Temecula1, M12, M23,

Ann-1, Dixon, GB514 e EB92-1), dos genomas descritos neste trabalho (cepas J1a12, Pr8x e Hib4) e de outros genomas que estão sendo sequenciados em nosso grupo (cepas U24d e Fb7, isoladas de citros, e cepa 3124, isolada de café) (Santana *et al.*, 2010). Como grupo externo foi utilizado o conjunto das sequências dos ortólogos de *Xanthomonas albilineans* (Pieretti *et al.*, 2009), espécie pertencente ao grupo das Xanthomonadaceae.

Para a reconstrução da árvore filogenética utilizamos o método probabilístico de Máxima Verossimilhança (ML), com o modelo geral de tempo-reverso (GTR +G) (Figura 23). Utilizando o grupo de genes *housekeeping* utilizado em cepas de *Xylella* (Nunney *et al.*, 2012), as cepas Hib4 e Pr8x posicionaram-se juntas em um clado basal separado das outras cepas sul-americanas e mais próximas das cepas norte-americanas. Este dado corrobora com o gráfico de “dot plot” entre as duas cepas (Figura 22), o qual mostrou extensa similaridade entre seus genomas. A cepa J1a12 posiciou-se separada, mas logo após o clado de outras cepas isoladas de citros (9a5c e U24d) (Figura 23).

Uma importante característica observada em filogenias de *X. fastidiosa* são os longos ramos basais levando a ramos terminais muito pequenos que correspondem a agrupamentos de hospedeiros específicos. Dessa forma, os dois maiores grupos, da América do Norte e do Sul, mostram grande diferenciação filogenética, porém pouca variação entre cepas do mesmo hospedeiro, o que sugere que as linhagens sofrem fortes pressões seletivas de seus hospedeiros vegetais (Schuenzel *et al.*, 2005; Doddapaneni *et al.*, 2007).

Além disso, outro estudo sugere a presença de eventos de recombinação entre isolados sul-americanos de *X. fastidiosa* e indica que tais eventos sejam mais frequentes do que entre isolados na América do Norte. Os resultados deste estudo também mostraram que a recombinação em *X. fastidiosa* pode estar ocorrendo não apenas dentro de subespécies, mas também entre elas (Almeida *et al.*, 2008). Assim, a possibilidade de introdução acidental de novos genótipos representa um risco real, uma vez que a mistura de subespécies de *Xylella* poderia gerar novas recombinações, e consequentemente gerar novas doenças e expandir a sua abrangência de hospedeiros. Precauções gerais devem ser tomadas contra a importação de subespécies não-nativas deste patógeno ao longo das Américas (Nunney *et al.*, 2011).

Devido a essa recombinação, não é possível inferir a história evolutiva desses *taxa*. Dessa forma, a concatenação de algumas sequências de genes tem como importância apenas no agrupamento de linhagens geneticamente e biologicamente

similares, mas limitado valor para inferência evolutiva (Almeida *et al.*, 2008). Portanto, estudos de filogenômica parecem ser mais adequados para inferência evolutiva de isolados de *Xylella fastidiosa*, uma vez que utilizam de um conjunto muito maior de sequências, fazendo com que o problema em utilizar possíveis sequências em processo de recombinação seja diminuído.

Figura 23. Árvore filogenética construída a partir de um grupo de sequências de 7 genes housekeeping. O método filogenético utilizado foi o de Máxima Verossimilhança, com modelo de substituição GTR + G. A árvore foi construída com as sequências dos genes *leuA* (XF1818), *petC* (XF0910), *malF* (XF2447), *cysG* (XF0832), *holC* (XF0136), *nuoL* (XF0316) e *gltT* (XF0656). As análises foram realizadas utilizando o software PAUP4b10 com 1000 réplicas (*bootstrap*). As árvores foram construídas com *Xanthomonas albilineans* como grupo externo. O número das ORFs entre parênteses refere-se ao locus na cepa 9a5c.

4.7. Ensaio de infecção da cepa J1a12 em plantas de cafeeiro

As relações entre isolados de *X. fastidiosa* de cafeeiros e citros no Brasil tem sido amplamente discutida e controversa. A CVC emergiu em 1987 no norte do Estado de São Paulo e a CLS foi pela primeira vez relatada em 1995 na mesma região. Entretanto, alguns autores têm argumentado que os sintomas de CLS estavam presentes em cafeeiros muito antes de 1987 e foram mal diagnosticados como infecções por nematóides ou desequilíbrio nutricional da planta. Além disso, *X. fastidiosa* foi posteriormente encontrada infectando cafeeiros em outras regiões do Brasil, sugerindo uma associação antiga com este hospedeiro (Almeida *et al.*, 2008). A alta frequência de detecção do patógeno em amostras de campo sugere que CLS é mais amplamente distribuída do que PLS ou CVC (Lopes *et al.*, 2003).

Historicamente, grandes produtores de café no Estado de São Paulo, com o declínio da indústria cafeeira, substituíram suas produções por extensos pomares de laranjeiras. Além disso, plantações de cafeeiros podem ser encontradas adjacentes a pomares de citros com CVC (Li *et al.*, 2001). Uma vez que a cultura de cafeeiros é anterior a citricultura, e que existe proximidade genética entre isolados de *X. fastidiosa* de café e citros, é possível que as linhagens de citros tenham se originado como uma variante clonal da população de linhagens de *X. fastidiosa* de cafeeiro (Qin *et al.*, 2001; Redak *et al.*, 2004; da Silva *et al.*, 2007). Além disso, linhagens de citros foram relatadas como patogênicas em cafeeiro em experimentos de inoculação cruzada (Li *et al.*, 2001). Em estudos filogenéticos, linhagens de *X. fastidiosa* isoladas de café e laranjeira doce de São Paulo não foram separadas em subgrupos independentes em relação a hospedeiros de origem, indicando proximidade genética (Montero-Astúa *et al.*, 2007). Análises filogenéticas preliminares indicaram uma proximidade filogenética entre as cepas J1a12 e 3124 (isolada de cafeeiro), que posteriormente não foi confirmada em novas análises a partir de um refinamento na montagem do genoma da cepa 3124. Baseado nestas análises preliminares, hipotetizamos que a cepa J1a12, isolada de citros e não-virulenta, poderia ter sua origem em plantas de cafeeiro, onde poderia, eventualmente, causar CLS (Escaldadura da folha de cafeeiro).

Para testar esta hipótese, realizamos infecções em cafeeiros com as cepas J1a12 e 3124. Para tal, inoculamos 24 plantas jovens com as cepas J1a12 ou 3124. Após 3 e 8 meses, as plantas foram visualmente inspecionadas para sintomas de CLS e a colonização de *X. fastidiosa* foi avaliada por PCR com o par de *primers* CVC-1/272-2-int (Pooler e Hartung, 1995), após extração de DNA do pecíolo de algumas

folhas (Murray e Thompson, 1980). Como controle negativo, uma planta foi inoculada apenas com tampão PBS e outra não foi inoculada.

Após 3 meses de infecção somente foi possível detectar DNA de *X. fastidiosa* em apenas duas plantas inoculadas com a cepa 3124 (dados não mostrados). Nenhuma das demais plantas após 3 ou 8 meses apresentou resultado positivo para presença da bactéria, indicando que ou a colonização não ocorreu ou o ensaio não foi tecnicamente adequado para sua detecção. A presença de inibidores de PCR tem sido um problema para a detecção de *X. fastidiosa* (Minsavage *et al.*, 1994; Bextine e Miller, 2004; Gambetta *et al.*, 2007). Em dois estudos independentes, o ensaio de PCR deu negativo para a presença de *Xylella*, mas foi positivo por ELISA, tanto na detecção de um isolado de chitalpa (Randall *et al.*, 2009) como de um isolado obtido de *Hibiscus syriacus* (McGaha *et al.*, 2007). Os autores sugerem que compostos da planta podem ter interferido no ensaio de PCR ou interagido com o anticorpo de *X. fastidiosa* gerando o resultado falso positivo no ensaio de ELISA (McGaha *et al.*, 2007). Tão pouco, nós não observamos sintomas de CLS nas plantas, exceto por uma única planta que apresentou, após 8 meses da inoculação com a cepa 3124, indícios do que pareciam ser dois dos sintomas de requeima do café (CLS), isto é surgimento de folhas menores e com menos pigmentação verde e diminuição da distância entre nós (Lima *et al.*, 1998) (dados não mostrados). As plantas inoculadas com a cepa J1a12 também não apresentaram sintomas aparentes da doença causada pela bactéria (dados não mostrados). As plantas inoculadas serão novamente inspecionadas dentro de mais alguns meses, visto que o aparecimento dos sintomas pode levar mais do que 12 meses (Dr. Helvécio Della Coletta-Filho, comunicação pessoal). Portanto, até o momento, nenhuma conclusão pode ser obtida em relação à colonização da cepa J1a12, isolada de laranjeira doce, em plantas de cafeeiro.

É importante mencionar que em um estudo anterior onde foram realizados experimentos de inoculação cruzada entre cepas de citros e de café, não se verificou, após 11 meses da inoculação, indícios de que cepas de citros infectam café ou vice-versa. Sintomas de CVC foram observados em plantas inoculadas com isolados de CVC enquanto que sintomas típicos de CLS não foram observados dentro do período de avaliação do experimento. Os autores sugeriram que os sintomas de CLS apareçam após anos de infecção em laboratório e que isolados de CVC podem se multiplicar, mas não sustentar infecção em plantas de café, enquanto que isolados de CLS não colonizam plantas de citros (Almeida *et al.*, 2008).

5. Conclusões

Neste trabalho descrevemos o pirosequenciamento e análises comparativas das sequências genômicas de três cepas de *X. fastidiosa* sendo uma delas isolada de citros e não-virulenta (J1a12), uma cepa isolada de ameixeira (Pr8x) e outra isolada de hibisco (Hib4).

A cepa J1a12 possui além de seu cromossomo principal de 2.788.789 pb dois plasmídeos, pXF51 e pXF27, respectivamente de 51.180 pb e 27.268 pb. pXF51 já foi descrito também na cepa de citros 9a5c (Simpson *et al.*, 2000; Marques *et al.*, 2001) e pXF27 tem similaridade com outros plasmídeos de cepas de *X. fastidiosa* norte-americanas isoladas de amoreira e videira. Demonstramos que pXF27 também está presente em outros isolados de diferentes localidades, principalmente no Estado de São Paulo. A análise comparativa da sequência genômica de J1a12 e 9a5c confirmou todas as diferenças anteriormente observadas utilizando-se microarranjos de DNA (Koide *et al.*, 2004), incluindo a ausência de uma cópia do gene da adesina da fimbria. Diferenças adicionais foram observadas tais como a ausência de 2 cópias das 3 CDS que codificam serina proteases da família de autotransportadores. Além disso, a cepa J1a12 parece ser deficiente na expressão de seu arsenal de toxinas protéicas e também no processamento de DSF (*Diffusible Signaling Factor*) que participa do sistema de sinalização célula-célula dependente da densidade celular, pois apresenta, respectivamente alterações em genes relacionados a produção de microcinas e no gene *rpfB*. Conjuntamente as diferenças observadas no genoma de J1a12 correlacionam-se com seu fenótipo menos agregativo e menos virulento.

A cepa Pr8x possui além de seu cromossomo principal de 2.666.242 pb um plasmídeo, pXF39, de 39.580 pb, o qual contém a maioria das CDS presentes no pXF51, além de CDS adicionais, como, por exemplo, do plasmídeo pXF5823 de outra cepa de *X. fastidiosa* de citros. A análise comparativa da sequência genômica de Pr8x e 9a5c mostraram a ausência de praticamente todo o sistema de conjugação inserido no cromossomo, o qual, provavelmente originou-se de pXF51. Assim como no genoma da cepa J1a12, também em Pr8x observamos a deleção da CDS da arginina deaminase (XF1250). Especulamos que uma CDS que codifica uma amido-hidrolase, e que está presente nestes genomas, poderia compensar a ausência da arginina deaminase que também hidrolisa ligações carbono-nitrogênio não peptídicas.

A cepa Hib4, isolada de hibisco, tem o maior cromossomo (2.813.297 pb) e também o maior plasmídeo (pXF64 com 64.251 pb) já descritos para *X. fastidiosa*.

pXF64 apresenta extensa similaridade com o plasmídeo pBVIE04 de *Burkholderia vietnamsis* cepa G4, sendo descrito pela primeira vez em cepas de *X. fastidiosa*. Entre as CDS ausentes no genoma de Hib4, destacamos a ortóloga de XF0077, que codifica um precursor de adesina de fimbrias e também está ausente na cepa J1a12. Além disso, a CDS ortóloga de XF1529, uma adesina não-fimbrial (Caserta *et al.*, 2010), também apresenta uma extensa deleção na sua porção central como verificado no genoma da cepa Pr8x. Se por um lado esta alteração parece não afetar o fenótipo agregativo da cepa Pr8x, podemos especular que ausência de duas adesinas possa estar correlacionada com o fenótipo de menor adesão observado no cultivo de Hib4.

É importante destacar que os genomas de Pr8x e Hib4 codificam a proteína precursora de poligalacturonase completa (ortóloga de XF2466 em 9a5c e PD1485 em Temecula). Nas cepas 9a5c e J1a12 e em outras isoladas de citros e de cafeeiros, este gene apresenta uma inserção que resulta no aparecimento de um códon de terminação prematuro, sendo que a ausência desta enzima tem sido associada a um fenótipo menos agressivo para estas cepas (Van Sluys *et al.*, 2003; Chatterjee *et al.*, 2008).

Finalmente, MLSA usando um grupo de genes *housekeeping*, cujas sequências foram recuperadas dos genomas de *X. fastidiosa* disponíveis, incluindo os descritos neste trabalho, mostraram, como esperado, a clara separação das cepas norte-americanas e sul-americanas. Porém, estudo anterior de reconstrução filogenética de linhagens sul-americanas indica que elas ainda estão em processo de diferenciação e recombinação, como sugerido anteriormente (Almeida *et al.*, 2008). A análise dos genomas das três cepas de *X. fastidiosa* mostrou a variedade e diversidade de regiões relacionadas a bacteriófagos e de plasmídeos que co-existem nas diferentes cepas deste fitopatógeno. As discrepâncias no conteúdo flexível do genoma são, de fato, o que mais diferencia as cepas de *X. fastidiosa*, resultado de eventos de aquisição de material genético por conjugação com outro organismo, por infecção por fagos ou mesmo por captação de moléculas de DNA do meio e subsequente incorporação no cromossomo.

6. Bibliografia

1. **Ahmadian, A., Ehn, M. e Hober, S.** 2006. Pyrosequencing: History, biochemistry and future. *Clinica Chimica Acta* 363:83-94.
2. **Almeida, R. P., Killiny, N., Newman, K. L., Chatterjee, S., Ionescu, M. e Lindow, S.** 2011. Contribution of *rpfB* to Cell-cell Signal Synthesis, Virulence, and Vector Transmission of *Xylella fastidiosa*. *Molecular plant-microbe interactions* : MPMI.
3. **Almeida, R. P. P., Blua, M. J., Lopes, J. R. S. e Purcell, A. H.** 2005. Vector transmission of *Xylella fastidiosa*: applying fundamental knowledge to generate disease management strategies. *Annals of Entomological Society of America* 98:775-786.
4. **Almeida, R. P. P., Nascimento, F. E., Chau, J., Prado, S. S., Tsai, C.-W., Lopes, S. A. e Lopes, J. R. S.** 2008. Genetic Structure and Biology of *Xylella fastidiosa* Strains Causing Disease in Citrus and Coffee in Brazil. *Applied and Environmental Microbiology* 74:3690-3701.
5. **Almeida, R. P. P. e Purcell, A. H.** 2006. Patterns of *Xylella fastidiosa* colonization on the precibarium of sharpshooter vectors relative to transmission to plants. *Annals of Entomological Society of America* 99:884-890.
6. **Altschul, S. F., Gish, W., Miller, W., Myers, E. W. e Lipman, D. J.** 1990. Basic Local Alignment Search Tool. *Journal of Molecular Biology* 215:403-410.
7. **Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. e Lipman, D. J.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25:3389-3402.
8. **Alves, E., Leite, B., Marucci, R. C., Pascholati, S. F., Lopes, J. R. S. e Andersen, P. C.** 2008. Retention sites for *Xylella fastidiosa* in four sharpshooter vectors (Hemiptera: Cicadellidae) analyzed by scanning electron microscopy. *Curr Microbiol* 56:531-538.
9. **Alves, E., Marucci, C. R., Lopes, J. R. S. e Leite, B.** 2004. Leaf Symptoms on Plum, Coffee and Citrus and the Relationship with the Extent of Xylem Vessels Colonized by *Xylella fastidiosa*. *Journal of Phytopathology* 152:291-297.
10. **Araújo, W. L., Marcon, J., Maccheroni Jr, W., van Elsas, J. D., van Vuurde, J. W. L. e Azevedo, J. L.** 2002. Diversity of endophytic bacterial populations and their interaction with *Xylella fastidiosa* in citrus plants. *Applied and Environmental Microbiology* 68:4906-4914.
11. **Asker, D. e Ohta, Y.** 2002. *Haloferax alexandrinus* sp nov., an extremely halophilic canthaxanthin-producing archaeon from a solar saltern in Alexandria (Egypt). *Int J Syst Evol Microbiol* 52:729-738.
12. **Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T. et al.** 2008. The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9:1-15.
13. **Backus, E. A. e Morgan, D. J. W.** 2011. Spatiotemporal colonization of *Xylella fastidiosa* in its vector supports the role of egestion in the inoculation mechanism of foregut-borne plant pathogens. *Phytopathology* 101:912-922.
14. **Bairoch, A.** 2000. The ENZYME database in 2000. *Nucleic Acids Research* 28:304-5.
15. **Belogurov, A. A., Delver, E. P., Agafonova, O. V., Belogurova, N. G., Lee, L. Y. e Kado, C. I.** 2000. Antirestriction protein ard (Type C) encoded by IncW plasmid pSa

- has a high similarity to the "protein transport" domain of TraC1 primase of promiscuous plasmid RP4. *Journal of Molecular Biology* 296:969-977.
16. **Berisha, B., Chen, Y. D., Zhang, G. Y., Xu, B. Y. e Chen, T. A.** 1998. Isolation of Pierce's disease bacteria from grapevine in Europe. *European Journal of Plant Pathology* 104:427-433.
 17. **Bextine, B., Lauzon, C., Potter, S., Lampe, D. e Miller, T. A.** 2004. Delivery of a genetically marked *Alcaligenes* sp. to the glassy-winged sharpshooter for use in a paratransgenic control strategy. *Curr Microbiol* 48:327-331.
 18. **Bextine, B. R. e Miller, T. A.** 2004. Comparison of whole-tissue and xylem fluid collection techniques to detect *Xylella fastidiosa* in grapevine and oleander. *Plant Disease* 88:600-604.
 19. **Bhattacharyya, A., Stilwagen, S., Ivanova, N., D'Souza, M., Bernal, A. et al.** 2002. Whole-genome comparative analysis of three phytopathogenic *Xylella fastidiosa* strains. *Proceedings of the National Academy of Sciences* 99:12403-12408.
 20. **Bove, J. M. e Ayres, A. J.** 2007. Etiology of three recent diseases of citrus in Sao Paulo State: sudden death, variegated chlorosis and huanglongbing. *IUBMB Life* 59:346-54.
 21. **Brlansky, R. H., Damsteegt, V. D. e Hartung, J. S.** 2002. Transmission of the citrus variegated chlorosis bacterium *Xylella fastidiosa* with the sharpshooter *Oncometopia nigricans*. *Plant Disease* 86:1237-1239.
 22. **Carver, T., Harris, S. R., Berriman, M., Parkhill, J. e McQuillan, J. A.** 2012. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* 28:464-9.
 23. **Caserta, R., Takita, M. A., Targon, M. L., Rosselli-Murai, L. K., de Souza, A. P., Peroni, L., Stach-Machado, D. R., Andrade, A., Labate, C. A., Kitajima, E. W., Machado, M. A. e de Souza, A. A.** 2010. Expression of *Xylella fastidiosa* fimbrial and afimbrial proteins during biofilm formation. *Applied and Environmental Microbiology* 76:4250-4259.
 24. **Ceglowski, P., Boitsov, A., Karamyan, N., Chai, S. e Alonso, J. C.** 1993. Characterization of the effectors required for stable inheritance of *Streptococcus pyogenes* pSM19035-derived plasmids in *Bacillus subtilis*. *Molecular and General Genetics* 241:579-585.
 25. **Chang, C. J., Garnier, M., Zreik, L., Rossetti, V. e Bove, J. M.** 1993. Culture and serological detection of the xylem-limited bacterium causing citrus variegated chlorosis and its identification as a strain of *Xylella fastidiosa*. *Current Microbiology* 27:137-142.
 26. **Chatelet, D. S., Wistrom, C. M., Purcell, A. H., Rost, T. L. e Matthews, M. A.** 2011. Xylem structure of four grape varieties and 12 alternative hosts to the xylem-limited bacterium *Xylella fastidiosa*. *annals of Botany*:1-13.
 27. **Chatterjee, S., Almeida, R. P. P. e Lindow, S.** 2008. Living in two worlds: the planta and insect lifestyle of *Xylella fastidiosa*. *Annual Review of Phytopathology* 46:243-271.
 28. **Chen, J. e Civerolo, E. L.** 2008. Morphological evidence for phages in *Xylella fastidiosa*. *Virology Journal* 5:1-4.

-
29. **Chen, J., Civerolo, E. L., Jarret, R. L., Van Sluys, M. A. e Oliveira, M. C.** 2005. Genetic discovery in *Xylella fastidiosa* through sequence analysis of selected randomly amplified polymorphic DNAs. *Curr Microbiol* 50:78-83.
 30. **Chen, J., Xie, G., Han, S., Chertkov, O., Sims, D. e Civerolo, E. L.** 2010. Whole Genome Sequences of Two *Xylella fastidiosa* Strains (M12 and M23) Causing Almond Leaf Scorch Disease in California. *Journal of Bacteriology* 192:4534.
 31. **Christie, P. J., Atmakuri, K., Krishnamoorthy, V., Jakubowski, S. e Cascales, E.** 2005. Biogenesis, Architecture, and Function of Bacterial Type IV Secretion Systems. *Annual Review of Microbiology* 59:451-485.
 32. **Ciraulo, M. B., Santos, D. S., Rodrigues, A. C. F. O., Oliveira, M. V., Rodrigues, T., Oliveira, R. C. e Nunes, L. R.** 2010. Transcriptome analysis of the phytopathogen *Xylella fastidiosa* growing under xylem-based chemical conditions. *Journal of Biomedicine and Biotechnology* 2010:1-18.
 33. **Coletta-Filho, H. D. e Machado, M. A.** 2003. Geographical genetic structure of *Xylella fastidiosa* from citrus in São Paulo State, Brazil. *Phytopathology* 93:28-34.
 34. **Coletta-Filho, H. D., Takita, M. A., de Souza, A. A., Aguilar-Vildoso, C. I. e Machado, M. A.** 2001. Differentiation of Strains of *Xylella fastidiosa* by a Variable Number of Tandem Repeat Analysis. *Applied and Environmental Microbiology* 67:4091-4095.
 35. **Cursino, L., Galvani, C. D., Athinuwat, D., Zaini, P. A., Li, Y., De La Fuente, L., Hoch, H. C., Burr, T. J. e Mowery, P.** 2011. Identification of an operon, Pil-Chp, that controls twitching motility and virulence in *Xylella fastidiosa*. *MPMI* 24:1198-1206.
 36. **Cussiol, J. R. R., Alves, S. V., Oliveira, M. A. e Netto, L. E. S.** 2003. Organic hydroperoxide resistance gene encodes a thiol-dependent peroxidase. *The Journal of Biological Chemistry* 278:11570-11578.
 37. **da Silva-Neto, J. F., Koide, T., Abe, C. M., Gomes, S. L. e Marques, M. V.** 2008. Role of σ^{54} in the regulation of genes involved in type I and type IV pili biogenesis in *Xylella fastidiosa*. *Arch Microbiol* 189:249-261.
 38. **da Silva, A. C., Ferro, J. A., Reinach, F. C., Farah, C. S., Furlan, L. R. et al.** 2002. Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature* 417:459-63.
 39. **da Silva, P. I. P.** 2010. Identificação de genes com expressão modulada por estreptomicina e de genes associados à virulência e patogenicidade em *Xylella fastidiosa*. Doutorado. Departamento de Bioquímica, Universidade de São Paulo, São Paulo. 138.
 40. **da Silva, V. S., Shida, C. S., Rodrigues, F. B., Ribeiro, D. C. D., de Souza, A. A., Coletta-Filho, H. D., Machado, M. A., Nunes, L. R. e de Oliveira, R. C.** 2007. Comparative genomic characterization of citrus-associated *Xylella fastidiosa* strains. *BMC Genomics* 8:1-15.
 41. **Damsteegt, V. D., Brlansky, R. H., Phillips, P. A. e Roy, A.** 2006. Transmission of *Xylella fastidiosa*, causal agent of citrus variegated chlorosis, by the glassy-winged sharpshooter, *Homalodisca coagulata*. *Plant Disease* 90:567-570.
 42. **Davis, M. J., French, W. J., Schaad, N. W.** 1981. Axenic culture of the bacteria associated with phony disease of peach and plum leaf scald. *Current Microbiology* 6:309-314.

-
43. **Davis, M. J., Purcell, A. H. e Thomson, S. V.** 1978. Pierce's disease of grapevines: isolation of the causal bacterium. *Science* 199:75-77.
 44. **Davis, M. W.** 2009. ApE- A Plasmid Editor. *In* ed. (ed.). Department of Biology - University of Utah, Utah.
 45. **de Mello Varani, A., Souza, R. C., Nakaya, H. I., de Lima, W. C., Paula de Almeida, L. G., Kitajima, E. W., Chen, J., Civerolo, E., Vasconcelos, A. T. e Van Sluys, M. A.** 2008. Origins of the *Xylella fastidiosa* prophage-like regions and their impact in genome differentiation. *PLoS ONE* 3:e4059.
 46. **de Souza, A. A., Takita, M. A., Coletta-Filho, H. D., Caldana, C., Goldman, G. H., Yanai, G. M., Muto, N. H., de Oliveira, R. C., Nunes, L. R. e Machado, M. A.** 2003. Analysis of gene expression in two growth states of *Xylella fastidiosa* and its relationship with pathogenicity. *Molecular Plant-Microbe Interactions* 16:867-75.
 47. **de Souza, A. A., Takita, M. A., Coletta-Filho, H. D., Caldana, C., Yanai, G. M., Muto, N. H., de Oliveira, R. C., Nunes, L. R. e Machado, M. A.** 2004. Gene expression profile of the plant pathogen *Xylella fastidiosa* during biofilm formation in vitro. *Fems Microbiology Letters* 237:341-53.
 48. **Delcher, A. L., Salzberg, S. L. e Phillippy, A. M.** 2003. Using MUMmer to identify similar regions in large sequence sets. *Current protocols in bioinformatics / editorial board, Andreas D. Baxevanis ... [et al.] Chapter 10:Unit 10 3.*
 49. **Doddapaneni, H., Francis, M., Yao, J., Lin, H. e Civerolo, E. L.** 2007. Genome-wide analysis of *Xylella fastidiosa*: implications for detection and strain relationships. *African Journal of Biotechnology* 6:055-066.
 50. **Dow, J. M. e Daniels, M. J.** 2000. *Xylella* genomics and bacterial pathogenicity to plants. *Yeast* 17:263-271.
 51. **Droege, M. e Hill, B.** 2008. The Genome Sequencer FLXTM System - Longer reads, more applications, straight forward bioinformatics and more complete data sets. *Journal of Biotechnology* 136:3-10.
 52. **Fivian-Hughes, A. S. e Davis, E. O.** 2010. Analyzing the regulatory role of the Higa antitoxin within *Mycobacterium tuberculosis*. *Journal of Bacteriology* 192:4348-4356.
 53. **Flemming, H. C. e Wingender, J.** 2010. The biofilm matrix. *Nature reviews. Microbiology* 8:623-33.
 54. **French, W. J. e Feliciano, A.** 1982. Distribution and severity of plum leaf scald disease in Brazil. *Plant Disease* 66:515-517.
 55. **French, W. J. e Kitajima, E. W.** 1978. Occurrence of plum leaf scald in Brazil and Paraguay. *Plant Disease Reporter* 62:1035-1038.
 56. **Fullwood, M. J., Wei, C.-L., Liu, E. T. e Ruan, Y.** 2009. Next-generation DNA sequencing of paired-end tags (PET) for transcriptome and genome analyses. . *Genome Research* 19:521-532.
 57. **Gambetta, G. A., Fei, J., Rost, T. L. e Matthews, M. A.** 2007. Leaf scorch symptoms are not correlated with bacterial populations during Pierce's disease. *Journal of Experimental Botany* 58:4037-4046.

-
58. **Garcia-Pino, A., Dao-Thi, M.-H., Gazit, E., Magnuson, R. D., Wyns, L. e Loris, R.** 2008. Crystallization of Doc and the Phd-Doc toxin-antitoxin complex Acta Crystallographica F64:1034-1038.
59. **Gogos, A., Mu, H., Bahna, F., Gomez, C. A. e Shapiro, L.** 2003. Crystal structure of YdcE protein from *Bacillus subtilis*. Proteins: Structure, Function, and Genetics 53:320-322.
60. **Gotfredsen, M. e Gerdes, K.** 1998. The *Escherichia coli relBE* genes belong to a new toxin-antitoxin gene family. Molecular Microbiology 29:1065-1076.
61. **Hagemann, M., Hasse, D. e Berg, G.** 2006. Detection of a phage genome carrying a zonula occludens like toxin gene (zot) in clinical isolates of *Stenotrophomonas maltophilia*. Archives of Microbiology 185:449-458.
62. **Hail, D., Lauzière, I., Dowd, S. E. e Bextine, B.** 2011. Culture independent survey of the microbiota of the glassy-winged sharpshooter (*Homalodisca vitripennis*) using 454 pyrosequencing. Environ. Entomol. 40:23-29.
63. **Hall, T.** 2005. BioEdit: Biological sequence alignment editor. In ed. (ed.). Ibis Therapeutics, Carlsbad, CA.
64. **Hanage, W. P., Fraser, C. e Spratt, B. G.** 2005. Fuzzy species among recombinogenic bacteria BMC Biology 3.
65. **Hernandez-Martinez, R., Cooksey, D. A. e Wong, F. P.** 2009. Leaf scorch of purple-leaved plum and sweetgum dieback: two new diseases in southern California caused by *Xylella fastidiosa* strains with different host ranges. Plant Disease 93:1131-1138.
66. **Hernandez-Martinez, R., Costa, H. S., Dumenyo, C. K. e Cooksey, D. A.** 2006. Differentiation of Strains of *Xylella fastidiosa* infecting grape, almonds, and oleander using a multiprimer PCR assay. Plant Disease 90:1382-1388.
67. **Hernandez-Martinez, R., de la Cerda, K. A., Costa, H. S., Cooksey, D. A. e Wong, F. P.** 2007. Phylogenetic relationships of *Xylella fastidiosa* strains isolated from landscape ornamentals in Southern California. Phytopathology 97:857-864.
68. **Hill, B. L. e Purcell, A. H.** 1995. Acquisition and retention of *Xylella fastidiosa* by and efficient vector, *Graphocephala atropunctata*. Phytopathology 85:209-212.
69. **Hill, B. L. e Purcell, A. H.** 1997. Populations of *Xylella fastidiosa* in plants required for transmission by an efficient vector. Phytopathology 87:1197-1201.
70. **Hok, S., Attard, A. e Keller, H.** 2010. Getting the most from the host: how pathogen force plants to cooperate in disease. MPMI 23:1253-1259.
71. **Hopkins, D. L.** 1989. *Xylella fastidiosa*: xylem-limited bacterial pathogen of plants. Annual Review of Phytopathology 27:271-90.
72. **Hopkins, D. L., Kirkpatrick, B. C., Hill, B. L., Smith, R. e Johnson, D.** 2010. Biological control of Pierce's disease of grapevine with benign strain of *Xylella fastidiosa* subsp. *piercei*
73. **Hopkins, D. L. e Purcell, A. H.** 2002. *Xylella fastidiosa*: Cause of Pierce's disease of grapevine and other emergent diseases. Plant Disease 86:1056-1066.

-
74. **Hunter, S., Jones, P., Mitchell, A., Apweiler, R., Attwood, T. K. et al.** 2012. InterPro in 2011: new developments in the family and domain prediction database. *Nucleic Acids Research* 40:D306-12.
75. **Hurley, J. M. e Woychik, N. A.** 2009. Bacterial toxin HigB associates with ribosomes and mediates translation-dependent mRNA cleavage at A-rich sites. . *The Journal of Biological Chemistry* 284:18605-18613.
76. **Huse, S. M., Huber, J. A., Morrison, H. G., Sogin, M. L. e Welch, D. M.** 2007. Accuracy and quality of massively parallel DNA pyrosequencing. *Genome biology* 8:R143.
77. **Igo, M. M., Huston, S. e Kirkpatrick, B. C.** 2010. The role of Type V secretion autotransporters in the virulence of *Xylella fastidiosa*. Presented at the 2010 Pierce's Disease Research Symposium, San Diego, California, USA.
78. **Janse, J. D. e Obradovic, A.** 2010. *Xylella fastidiosa*: its biology, diagnosis, control and risks. *Journal of Plant Pathology* 92:S1.35- S1.48.
79. **Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. e Tanabe, M.** 2012. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Research* 40:D109-14.
80. **Killiny, N. e Almeida, R. P. P.** 2009. Host structural carbohydrate induces vector transmission of a bacterial plant pathogen. *PNAS* 106:22416-22420.
81. **Killiny, N., Prado, S. S. e Almeida, R. P. P.** 2010. Chitin utilization by the insect-transmitted bacterium *Xylella fastidiosa*. *Applied and Environmental Microbiology* 76:6134-6140.
82. **Kitajima, E. W., Bakarcic, M. e Fernandez-Valiela, M. V.** 1975. Association of Rickettsialike Bacteria With Plum Leaf Scald Disease. *Phytopathology* 65:476-479.
83. **Kitajima, E. W., Coletta-Filho, H. D., Machado, M. A. e Novaes, Q. S.** 2000. Escaldadura das Folhas em *Hibiscus shizopetalus* associada à infecção por *Xylella fastidiosa* em Brasília, DF., p. 323, *Fitopatologia Brasileira*, vol. 25. XXXIII Congresso Brasileiro de Fitopatologia, Belém.
84. **Koide, T., Zaini, P. A., Moreira, L. M., Vencio, R. Z. N., Matsukuma, A. Y., Durham, A. M., Teixeira, D. C., El-Dorry, H., Monteiro, P. B., da Silva, A. C. R., Verjovski-Almeida, S., da Silva, A. M. e 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. . *Journal of Bacteriology* 186:5442-5449.
85. **Kung, S. H. e Almeida, R. P. P.** 2011. Natural Competence and Recombination in the Plant Pathogen *Xylella fastidiosa*. *Applied and Environmental Microbiology* 77:5278-5282.
86. **Lacava, P. T., Araújo, W. L., Marcon, J., Maccheroni Jr, W. e Azevedo, J. L.** 2004. Interaction between endophytic bacteria from citrus plants and the phytopathogenic bacteria *Xylella fastidiosa*, causal agent of citrus-variegated chlorosis. *Letters in Applied Microbiology* 39:55-59.
87. **Lacava, P. T., Parker, J., Andreote, F. D., Dini-Andreote, F., Ramirez, J. L. e Miller, T. A.** 2007. Analysis of the bacterial community in glassy-winged sharpshooter heads. *Entomological research* 37:261-266.

-
88. **Lambais, M. R., Goldman, M. H. S., Camargo, L. E. A. e Goldman, G. H.** 2000. A genomic approach to the understanding of *Xylella fastidiosa* pathogenicity. *Current Opinion in Microbiology* 3:459-462.
 89. **Latham, A. J. e Norton, J. D.** 1980. Incidence of Plum Leaf Scald in Alabama. *Agricultural Experiment Station/Auburn University* 585:3-15.
 90. **Lauzière, I. e Sétamou, M.** 2009. Suitability of Different Host Plants for Oviposition and Development of *Homalodisca vitripennis* (Hemiptera: Cicadellidae) and Its Implication on Mass Rearing. *Annals of Entomological Society of America* 102:642-649.
 91. **Lee, M. W., Rogers, E. E. e Stenger, D. C.** 2010. Functional Characterization of Replication and Stability Factors of an Incompatibility Group P-1 Plasmid from *Xylella fastidiosa*. *Applied and Environmental Microbiology* 76:7734-7740.
 92. **Leite, B., Ishida, M. L., Alves, E., Carrer, H., Pascholati, S. F. e Kitajima, E. W.** 2002. Genomics and X-ray microanalysis indicate that Ca^{2+} and thiols mediate the aggregation and adhesion of *Xylella fastidiosa*. *Brazilian Journal of Medical and Biological Research* 35:645-650.
 93. **Li, W.-B., Pria, W. D., Lacava, P. M., Qin, X. e Hartung, J. S.** 2003. Presence of *Xylella fastidiosa* in sweet orange fruit and seeds and its transmission to seedlings. *Phytopathology* 93:953-958.
 94. **Li, W.-B., Pria, W. D., Teixeira, D. C., Miranda, V. S., Ayres, A. J., Franco, C. F., Costa, M. G., He, C.-X., Costa, P. I. e Hartung, J. S.** 2001. Coffee Leaf Scorch Caused by a Strain of *Xylella fastidiosa* from Citrus. *Plant Disease* 85:501-505.
 95. **Li, W. B., Zreik, L., Fernandes, N. G., Miranda, V. S., Teixeira, D. C., Ayres, A. J., Garnier, M. e Bové, J. M.** 1999. A triply cloned strain of *Xylella fastidiosa* multiplies and induces symptoms of citrus variegated chlorosis in sweet orange. *Curr Microbiol* 39:106-108.
 96. **Lima, J. E. O., Miranda, V. S., Hartung, J. S., Bransky, R. H., Coutinho, A., Roberto, S. R. e Carlos, E. F.** 1998. Coffee Leaf Scorch Bacterium: Axenic Culture, Pathogenicity, and Comparison with *Xylella fastidiosa* of Citrus. *Plant Disease* 82:94-97.
 97. **Lopes, S. A., Marcussi, S., Torres, S. C. Z., Souza, V., Fagan, C., França, S. C., Fernandes, N. G. e Lopes, J. R. S.** 2003. Weeds as alternative hosts of the citrus, coffee, and plum strains of *Xylella fastidiosa* in Brazil. *Plant Disease* 87:544-549.
 98. **Ludwig, W.** 2007. Nucleic acid techniques in bacterial systematics and identification. *International Journal of Food Microbiology* 120:225-236.
 99. **Maiden, M. C. J., Bygraves, J. A., Feil, E., Morelli, G., Russel, J. E., Urwin, R., Zhang, Q., Zhou, J. J., Zurth, K., Caugant, D. A., Feavers, I. M., Achtman, M. e Spratt, B. G.** 1998. Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. *Proceedings of the National Academy of Sciences of the United States of America* 95:3140-3145.
 100. **Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S. et al.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437:376-380.
 101. **Markowitz, V. M., Chen, I. M., Palaniappan, K., Chu, K., Szeto, E., Grechkin, Y., Ratner, A., Jacob, B., Huang, J., Williams, P., Huntemann, M., Anderson, I.,**

- Mavromatis, K., Ivanova, N. N. e Kyrpides, N. C.** 2012. IMG: the Integrated Microbial Genomes database and comparative analysis system. *Nucleic Acids Research* 40:D115-22.
102. **Marques, M. V., da Silva, A. M. e Gomes, S. L.** 2001. Genetic organization of plasmid pXF51 from the plant pathogen *Xylella fastidiosa*. *Plasmid* 45:184-199.
103. **Matsumoto, A. e Igo, M. M.** 2010. Species-specific type II restriction-modification system of *Xylella fastidiosa* temecula1. *Applied and Environmental Microbiology* 76:4092-5.
104. **Mattick, J. S.** 2002. Type IV pili and twitching motility. *Annual Review of Microbiology* 56:289-314.
105. **McGaha, L. A., Jackson, B., Bextine, B., McCullough, D. e Morano, L.** 2007. Potential Plant Reservoirs for *Xylella fastidiosa* in South Texas. *American Journal of Enology and Viticulture* 58:398-401.
106. **McTaggart, L. R., Richardson, S. E., Witkowska, M. e Zhang, S. X.** 2010. Phylogeny and Identification of *Nocardia* Species on the Basis of Multilocus Sequence Analysis. *Journal of Clinical Microbiology* 48:4525-4533.
107. **Mehta, A. e Rosato, Y. B.** 2001. Phylogenetic relationships of *Xylella fastidiosa* strains from different hosts, based on 16S rDNA and 16S-23S intergenic spacer sequences. *Int J Syst Evol Microbiol* 51:311-318.
108. **Meng, Y., Li, Y., Galvani, C. D., Hao, G., Turner, J. N., Burr, T. J. e Hoch, H. C.** 2005. Upstream migration of *Xylella fastidiosa* via pilus-driven twitching motility. *Journal of Bacteriology* 187:5560-5567.
109. **Meyer, M. M.** 2010. The effects of cold temperature on *Xylella fastidiosa*-infected *Vitis vinifera*. Plant Pathology, University of California, Davis.
110. **Minsavage, G. V., Thompson, C. M., Hopkins, D. L., Leite, R. M. V. B. C. e Stall, R. E.** 1994. Development of a polymerase chain reaction protocol for detection of *Xylella fastidiosa* in plant tissue. *Phytopathology* 84:456-461.
111. **Molina, R. O., Gonçalves, A. M. O., Zanutto, A. e Nunes, W. M. C.** 2010. Populational fluctuation of vectors of *Xylella fastidiosa*, Wells in sweet orange [*Citrus sinensis* (L.) Osbeck] varieties of Northwest Paraná State, Brazil. *Brazilian Archives of Biology and Technology* 53:549-554.
112. **Monteiro, P. B., Teixeira, D. C., Palma, R. R., Garnier, M., Bove, J. M. e Renaudin, J.** 2001. Stable transformation of the *Xylella fastidiosa* Citrus Variegated Chlorosis Strain with *oriC* Plasmids. *Applied and Environmental Microbiology* 67:2263-2269.
113. **Montero-Astúa, M., Hartung, J. S., Aguilar, E., Chacón, C., Li, W., Albertazzi, F. J. e Rivera, C.** 2007. Genetic Diversity of *Xylella fastidiosa* strains from Costa Rica, São Paulo, Brazil, and United States. *Phytopathology* 97:1338-1347.
114. **Moreira, L. M., de Souza, R. F., Almeida Jr., N. F., Setubal, J. C., Oliveira, J. C. F., Furlan, L. R., Ferro, J. A. e da Silva, A. C. R.** 2004. Comparative genomics analyses of citrus-associated bacteria. *Annu. Rev. Phytopathol.* 42:163-184.
115. **Moreira, L. M., De Souza, R. F., Digiampietri, L. A., Da Silva, A. C. e Setubal, J. C.** 2005. Comparative analyses of *Xanthomonas* and *Xylella* complete genomes. *Omics* 9:43-76.

-
116. **Muller, C.** 2009. Análise faunística e flutuação populacional de cigarrinhas (Hemiptera: Cicadellidae) potenciais vetoras de *Xylella fastidiosa* em pomares de ameixeira nos estados do Rio Grande do Sul e São Paulo, Brasil. Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, São Paulo. 68.
117. **Murray, M. G. e Thompson, W. F.** 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* 8:4321-4325.
118. **Murugaiyan, S., Bae, J. Y., Wu, J., Lee, S. D., Um, H. Y., Choi, H. K., Chung, E., Lee, J. H. e Lee, S. W.** 2011. Characterization of filamentous bacteriophage PE226 infecting *Ralstonia solanacearum* strains. *Journal of Applied Microbiology* 110:296-303.
119. **Mutschler, H., Gebhardt, M., Shoeman, R. L. e Meinhart, A.** 2011. A novel mechanism of programmed cell death in bacteria by toxin-antitoxin systems corrupts peptidoglycan synthesis. *PLoS Biol* 9:1-12.
120. **Mutschler, H. e Meinhart, A.** 2011. ϵ/ζ systems: their role in resistance, virulence, and their potencial for antibiotic development. *J Mol Med*.
121. **Nadarasah, G. e Stavrinides, J.** 2011. Insects as alternative hosts for phytopathogen bacteria. *FEMS Microbiol Rev* 35:555-575.
122. **Neves, M. F., Lopes, F. F., Trombin, V. G., Amaro, A. A., Neves, E. M. e Jank, M. S.** 2007. Caminhos para a citricultura: uma agenda para manter a liderança mundial. Atlas.
123. **Neves, M. F., Trombin, V. G., Milan, P., Lopes, F. F., Cressoni, F. e Kalaki, R.** 2011. O retrato da citricultura brasileira. In M. F. Neves (ed.). FEA-RP/ FUNDACE, Ribeirão Preto.
124. **Nunes, L. R., Rosato, Y. B., Muto, N. H., Yanai, G. M., da Silva, V. S., Leite, D. B., Gonçalves, E. R., Souza, A. A., Coletta-Filho, H. D., Machado, M. A., Lopes, S. A. e Oliveira, R. C.** 2003. Microarray analyses of *Xylella fastidiosa* provide evidence of coordinated transcription control of laterally transferred elements. *Genome Research* 13:570-578.
125. **Nunney, L., Elfekih, S. e Stouthamer, R.** 2012. The importance of multilocus sequence typing: cautionary tales from the bacterium *Xylella fastidiosa*. *Phytopathology*:1-23.
126. **Nunney, L., Yuan, X., Bromley, R., Hartung, J. S., Montero-Astúa, M., Moreira, L., Ortiz, B. e Stouthamer, R.** 2011. Population genomic analysis of a bacterial plant pathogen: novel insight into the origin of Pierce's disease of grapevine in the U.S. *PLoS ONE* 5:1-9.
127. **Pachulec, E. e van der Does, C.** 2010. Conjugative plasmids of *Neisseria gonorrhoeae*. *PLoS One* 5:1-14.
128. **Pashalidis, S., Moreira, L. M., Zaini, P. A., Campanharo, J. C., Alves, L. M., Ciapina, L. P., Vencio, R. Z., Lemos, E. G., Da Silva, A. M. e Da Silva, A. C.** 2005. Whole-genome expression profiling of *Xylella fastidiosa* in response to growth on glucose. *Omics : a journal of integrative biology* 9:77-90.
129. **Pérez-Donoso, A. G., Greve, L. C., Walton, H., Shackel, K. A. e Labavitch, J. M.** 2007. *Xylella fastidiosa* infection and ethylene exposure result in xylem and water movement disruption in grapevine shoots. *Plant Physiology* 143:1024-1036.

-
130. **Pérez-Donoso, A. G., Sun, Q., Roper, M. C., Greve, L. C., Kirkpatrick, B. C. e Labavitch, J. M.** 2010. Cell Wall-Degrading Enzymes Enlarge the Pore Size of Intervessel Pit Membranes in Healthy and *Xylella fastidiosa*-Infected Grapevines. *Plant Physiology* 152:1748-1759.
131. **Pieretti, I., Royer, M., Barbe, V., Carrere, S., Koebnik, R., Cociancich, S., Couloux, A., Darrasse, A., Gouzy, J., Jacques, M. A., Lauber, E., Manceau, C., Mangenot, S., Poussier, S., Segurens, B., Szurek, B., Verdier, V., Arlat, M. e Rott, P.** 2009. The complete genome sequence of *Xanthomonas albilineans* provides new insights into the reductive genome evolution of the xylem-limited Xanthomonadaceae. *BMC Genomics* 10:616.
132. **Pinel, N., Davidson, S. K. e Stahl, D. A.** 2008. *Verminephrobacter eiseniae* gen. nov., sp. nov., a nephridial symbiont of the earthworm *Eisenia foetida* (Savigny). *International Journal of Systematic and Evolutionary Microbiology* 58:2147-2157.
133. **Pooler, M. R. e Hartung, J. S.** 1995. Specific PCR detection and identification of *Xylella fastidiosa* strains causing citrus variegated chlorosis. *Current Microbiology* 31:377-381.
134. **Punta, M., Coggill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., Heger, A., Holm, L., Sonnhammer, E. L., Eddy, S. R., Bateman, A. e Finn, R. D.** 2012. The Pfam protein families database. *Nucleic Acids Research* 40:D290-301.
135. **Purcell, A. H. e Finlay, A.** 1979. Evidence for noncirculative transmission of Pierce's disease bacterium by sharpshooter leafhoppers. *Phytopathology* 69:393-395.
136. **Purcell, A. H. e Hopkins, D. L.** 1996. Fastidious xylem-limited bacterial plant pathogens. *Annu. Rev. Phytopathol.* 34:131-151.
137. **Qin, X. e Hartung, J. S.** 2001. Construction of a Shuttle Vector and Transformation of *Xylella fastidiosa* with Plasmid DNA. *Curr Microbiol* 43:158-162.
138. **Qin, X., Miranda, V. S., Machado, M. A., Lemos, E. G. M. e Hartung, J. S.** 2001. An evaluation of the genetic diversity of *Xylella fastidiosa* isolated from diseased citrus and coffee in São Paulo, Brazil. *Phytopathology* 91:599-605.
139. **Raju, B. C., Wells, J. M., Nyland, G., Brlansky, R. H. e Lowe, S. K.** 1982. Plum Leaf Scald: Isolation, Culture, and Pathogenicity of the Causal Agent. *Phytopathology* 72:1460-1466.
140. **Randall, J. J., Goldberg, N. P., Kemp, J. D., Radionenko, M., French, J. M., Olsen, M. W. e Hanson, S. F.** 2009. Genetic Analysis of a Novel *Xylella fastidiosa* Subspecies Found in the Southwestern United States. *Applied and Environmental Microbiology* 75:5631-5638.
141. **Redak, R. A., Purcell, A. H., Lopes, J. R. S., Blua, M. J., Mizell III, R. F. e Andersen, P. C.** 2004. The biology of xylem fluid-feeding insect vectors of *Xylella fastidiosa* and their relation to disease epidemiology. *Annual Review of Entomology* 49:243-70.
142. **Rodrigues, J. L. M., Silva-Stenico, M. E., Gomes, J. E., Lopes, J. R. S. e Tsai, S. M.** 2003. Detection and Diversity Assessment of *Xylella fastidiosa* in Field-Collected Plant and Insect Samples by Using 16S rRNA and gyrB Sequences. *Applied and Environmental Microbiology* 69:4249-4255.
143. **Rooney, A. P., Price, N. P. J., Ehrhardt, C., Swezey, J. L. e Bannan, J. D.** 2009. Phylogeny and molecular taxonomy of the *Bacillus subtilis* species complex and

- description of *Bacillus subtilis* subsp *inaquosorum* subsp nov. Int J Syst Evol Microbiol 59:2429-2436.
144. **Roper, M. C., Greve, L. C., Warren, J. G., Labavitch, J. M. e Kirkpatrick, B. C.** 2007. *Xylella fastidiosa* requires polygalacturonase for colonization and pathogenicity in *Vitis vinifera* grapevines. MPMI 20:411-419.
145. **Rossetti, V. e De Negri, D.** 1990. Clorose Variegada dos Citros: revisão. Laranja 11:1-14.
146. **Rossetti, V., Garnier, M., Bove, J. M., Beretta, M. J. G., Teixeira, A. R. R., Quaggio, J. A. e Denegri, J. D.** 1990. Occurrence of xylem-restricted bacteria in sweet orange trees affected by chlorotic variegation, a new citrus disease in Brazil. Comptes Rendus de L Academie Des Sciences 310:345-349.
147. **Rozen, S. e Skaletsky, H.** 2000. Primer3 on the WWW for general users and for biologist programmers. Methods in Molecular Biology 132:365-386.
148. **Rudrappa, T., Biedrzycki, M. L. e Bais, H. P.** 2008. Causes and consequences of plant-associated biofilms. Fems Microbiology Ecology 64:153-166.
149. **Ryan, R. P. e Dow, J. M.** 2011. Communication with a growing family: diffusible signal factor (DSF) signaling in bacteria. Trends in microbiology 19:145-52.
150. **Saillard, C., Carle, P., Duret-Nurbel, S., Henri, R., Killiny, N., Carrère, S., Gouzy, J., Bové, J.-M., Renaudin, J. e Foissac, X.** 2008. The abundant extrachromosomal DNA content of the *Spiroplasma citri* GII3-3X genome. BMC Genomics 9:1-13.
151. **Santana, W. O., Beckedorff, F. C. F., Amaral, M. S., Coletta-Filho, H. D., Souza, A. A., Machado, M. A., Almeida, L. G. P., Vasconcelos, A. T. R., Verjovski-Almeida, S. e da Silva, A. M.** 2010. Pirosequenciamento, genômica comparativa e filogenia de cepas de *Xylella fastidiosa*. Presented at the 56° Congresso Brasileiro de Genética, Casa Grande Hotel Resort, Guarujá, SP, Brasil.
152. **Schaad, N. W., Postnikova, E., Lacy, G., Fatmi, M. e Chang, C.-J.** 2004. *Xylella fastidiosa* subspecies: *Xylella fastidiosa* subsp. *piercei*, subsp. nov., *Xylella fastidiosa* subsp. *multiplex* subsp. nov., and *X. fastidiosa* subsp. *pauca* subsp. nov. System. Appl. Microbiol. 27:290-300.
153. **Schreiber, H. L. I., Koirala, M., Lara, A., Ojeda, M., Dowd, S. E., Bextine, B. e Morano, L.** 2010. Unraveling the first *Xylella fastidiosa* subsp. *fastidiosa* genome from Texas. Southwestern Entomologist 35:479-483.
154. **Schuenzel, E. L., Scally, M., Stouthamer, R. e Nunney, L.** 2005. A Multigene Phylogenetic Study of Clonal Diversity and Divergence in North American Strains of the Plant Pathogen *Xylella fastidiosa*. Applied and Environmental Microbiology 71:3832-3839.
155. **Selengut, J. D., Haft, D. H., Davidsen, T., Ganapathy, A., Gwinn-Giglio, M., Nelson, W. C., Richter, A. R. e White, O.** 2007. TIGRFAMs and Genome Properties: tools for the assignment of molecular function and biological process in prokaryotic genomes. Nucleic Acids Research 35:D260-4.
156. **Seshasayee, A. S., Fraser, G. M. e Luscombe, N. M.** 2010. Comparative genomics of cyclic-di-GMP signalling in bacteria: post-translational regulation and catalytic activity. Nucleic Acids Research 38:5970-81.
157. **Shi, X. Y., Dumenyo, C. K., Hernandez-Martinez, R., Azad, H. e Cooksey, D. A.** 2007. Characterization of regulatory pathways in *Xylella fastidiosa*: genes and

- phenotypes controlled by *algU*. Applied and Environmental Microbiology 73:6748-6756.
158. **Shi, X. Y., Dumenyo, C. K., Hernandez-Martinez, R., Azad, H. e Cooksey, D. A.** 2009. Characterization of regulatory pathways in *Xylella fastidiosa*: genes and phenotypes controlled by *gacA*. Applied and Environmental Microbiology 75:2275-2283.
159. **Shneyer, V. S.** 2007. On the species-specificity of DNA: Fifty years later. Biochemistry-Moscow 72:1377-1384.
160. **Simpson, A. J., Reinach, F. C., Arruda, P., Abreu, F. A., Acencio, M. et al.** 2000. The genome sequence of the plant pathogen *Xylella fastidiosa*. Nature 406:151-159.
161. **Smith, C. L.** 1991. Separation and analysis of DNA by electrophoresis. Current Opinion in Biotechnology 2:86-91.
162. **Stenger, D. C., Lee, M. W., Rogers, E. E. e Chen, J.** 2010. Plasmids of *Xylella fastidiosa* mulberry-infecting strains share extensive sequence identity and gene complement with pVEIS01 from the earthworm symbiont *Verminephrobacter eiseniae*. Physiological and Molecular Plant Pathology 74:238-245.
163. **Studholme, D. J., Downie, J. A. e Preston, G. M.** 2005. Protein domains and architectural innovation in plant-associated Proteobacteria. BMC Genomics 6:17.
164. **Summer, E. J., Enderle, C. J., Ahern, S. J., Gill, J. J., Torres, C. P., Appel, D. N., Black, M. C., Young, R. e Gonzalez, C. F.** 2010. Genomic and biological analysis of phage Xfas53 and related prophages of *Xylella fastidiosa*. Journal of Bacteriology 192:179-190.
165. **Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V., Krylov, D. M., Mazumder, R., Mekhedov, S. L., Nikolskaya, A. N., Rao, B. S., Smirnov, S., Sverdlov, A. V., Vasudevan, S., Wolf, Y. I., Yin, J. J. e Natale, D. A.** 2003. The COG database: an updated version includes eukaryotes. BMC bioinformatics 4:41.
166. **Thompson, J. D., Gibson, T. J. e Higgins, D. G.** 2002. Multiple sequence alignment using ClustalW and ClustalX. Current protocols in bioinformatics / editorial board, Andreas D. Baxevanis ... [et al.] Chapter 2:Unit 2 3.
167. **Thorne, E. T., Stevenson, J. F., Rost, T. L., Labavitch, J. M. e Matthews, M. A.** 2006. Pierce's disease symptoms: comparison with symptoms of water deficit and the impact of water deficits. Am. J. Enol. Vitic. 57:1-11.
168. **Tock, M. R. e Dryden, D. T.** 2005. The biology of restriction and anti-restriction. Current Opinion in Microbiology 8:466-72.
169. **Toledo, M. A. S., Schneider, D. R., Azzoni, A. R., Favaro, M. T. P., Pelloso, A. C., Santos, C. A., Saraiva, A. M. e Souza, A. P.** 2011. Characterization of an oxidative stress response regulator, homologous to *Escherichia coli* OxyR, from the phytopathogen *Xylella fastidiosa*. Protein Expression and Purification 75:204-210.
170. **Van Melderren, L.** 2010. Toxin-antitoxin systems: why so many, what for? Current Opinion in Microbiology 13:1-5.
171. **Van Melderren, L. e De Bast, M. S.** 2009. Bacterial Toxin-Antitoxin Systems: More than Selfish Entities? PLoS Genetics 5:1-6.

-
172. **Van Sluys, M. A., de Oliveira, M. C., Monteiro-Vitorello, C. B., Miyaki, C. Y., Furlan, L. R. et al.** 2003. Comparative Analyses of the Complete Genome Sequences of Pierce's Disease and Citrus Variegated Chlorosis Strains of *Xylella fastidiosa*. *Journal of Bacteriology* 185:1018-1026.
173. **Voegel, T. M., Warren, J. G., Matsumoto, A., Igo, M. M. e Kirkpatrick, B. C.** 2010. Localization and characterization of *Xylella fastidiosa* haemagglutinin adhesins. *Microbiology* 156:2172-2179.
174. **Wells, J. M., Raju, B. C., Hung, H.-Y., Weisburg, W. G., Mandelco-Paul, L. e Brenner, D. J.** 1987. *Xylella fastidiosa* gen. nov., sp. nov: Gram-Negative, Xylem-Limited, Fastidious Plant Bacteria Related to *Xanthomonas* spp. . *International Journal of Systematic Bacteriology* 37:136-143.
175. **Wells, J. M., Raju, B. C., Nyland, G. e Lowe, S. K.** 1981. Medium for Isolation and Growth of Bacteria Associated with Plum Leaf Scald and Phony Peach Diseases. *Applied and Environmental Microbiology* 42:357-363.
176. **Wilgenbusch, J. C. e Swofford, D.** 2003. Inferring evolutionary trees with PAUP*. *Curr Protoc Bioinformatics* 6.
177. **Williams, K. P., Gillespie, J. J., Sobral, B. W., Nordberg, E. K., Snyder, E. E., Shallom, J. M. e Dickerman, A. W.** 2010. Phylogeny of gammaproteobacteria. *Journal of Bacteriology* 192:2305-14.
178. **Xu, Y., Lee, M., Moeller, A., Song, S., Yoon, B. Y., Kim, H. M., Jun, S. Y., Lee, K. e Ha, N. C.** 2011. Funnel-like hexameric assembly of the periplasmic adapter protein in the tripartite multidrug efflux pump in gram-negative bacteria. *The Journal of Biological Chemistry* 286:17910-20.
179. **Yang, L., Hong, L., Takahashi, Y., Chen, F., Walker, M. A. e Civerolo, E. L.** 2011. Proteomic analysis of grapevine stem in response to *Xylella fastidiosa* inoculation. *Physiological and Molecular Plant Pathology* 75:90-99.
180. **Zaini, P. A., De La Fuente, L., Hoch, H. C. e Burr, T. J.** 2009. Grapevine xylem sap enhances biofilm development by *Xylella fastidiosa*. *FEMS Microbiol Lett* 295:129-134.
181. **Zaini, P. A., Fogaca, A. C., Lupo, F. G., Nakaya, H. I., Vencio, R. Z. e da Silva, A. M.** 2008. The iron stimulon of *Xylella fastidiosa* includes genes for type IV pilus and colicin V-like bacteriocins. *Journal of Bacteriology* 190:2368-78.
182. **Zambolim, L. e Bassanezi, R. B.** 2006. Doenças quarentenárias dos citros. Universidade Federal de Viçosa.
183. **Zhang, S., Flores-Cruz, Z., Kumar, D., Chakrabarty, P., Hopkins, D. L. e Gabriel, D. W.** 2011. The *Xylella fastidiosa* Biocontrol Strain EB92-1 Genome is Very Similar and Syntenic to Pierce's Disease Strains. *Journal of Bacteriology* 193:5576-5577.

7. Apêndice

Figura S1. Alinhamento da sequência de nucleotídeos (A) e de aminoácidos (B) da poligalacturonase (pglA) das cepas 9a5c, Temecula, Pr8x e Hib4. A seta aponta a inserção que resulta em mutação de *frameshift* na cepa 9a5c.

A)

pglA_9a5c	ATGAACCTTGACCGTTTCCTCCCCCTAGTCTTTGGACTGCACTGTTTGGCCGAATGGGC	60
pglA_PR8X	ATGAACCTTGACCGTTTCCTCCCCCTAGTCTTTGGACTGCACTGTTTGGCCGAATGGGC	60
pglA_Hib4	ATGAACCTTGACCGTTTCCTCCCCCTAGTCTTTGGACTGCACTGTTTGGCCGAATGGGC	60
pglA_Tem	-----GTGTTTGGACTGCACTGTTTGGCCGAATGGGC	33

pglA_9a5c	GGCACTGCT-----GATGTCTTTCGAGGATCACAGCCTTCCCTATCCGATCTTTCC	111
pglA_PR8X	GGCACTGCT-----GATGTCTTTCGAGGATCACAGCCTTCCCTATCCGATCTTTCC	111
pglA_Hib4	GGCACTGCTGATACTGTCTTTCGAGTGTACAGCGTCCCATCCGATCTTTCC	120
pglA_Tem	GGCACTGCT-----GATGTCTTTCGAGTGTACAGCTGTCTCATCAGATCTTTCC	84

pglA_9a5c	GAGCCGGTATCGGTGAAAACGCTCTGGGGCGAGGTGCAACGTCCCAGCCTACCTACCCAC	171
pglA_PR8X	GAGCCGGTATCGGTGAAAACGCTCTGGGGCGAGGTGCAACGTCCCAGCCTACCTACCCAC	171
pglA_Hib4	GAGCCGGTATCGGTGAAAACGCTCTGGGGCGAGGTGCAACGTCCCAGCCTACCTACCCAC	180
pglA_Tem	GAGCCGGTATCGGTGAAAACGCTATGGGGCGAGGTGCAACGTCCCAGCCTACCTACCCAC	144

pglA_9a5c	GTGTGTACGGTATTGCCTGCGCGCCTGACTCCTGAACATGGCTCCCTTGACGCACTCGAT	231
pglA_PR8X	GTGTGTACGGTATTGCCTGCGCGCCTGACTCCTGAACATGGCTCCCTTGACGCACTCGAT	231
pglA_Hib4	GTGTGTACGGTATTGCCTGCGCGCCTGACTCCTGAACATGACTCCCTTGACGCACTTGAT	240
pglA_Tem	GTGTGTACGGTATTGCCTGCGCGCCTGACTCCTAAACATGGCTCCATTGACGCACTCGAT	204

pglA_9a5c	GCCAATCCACGGGCTCCAAGCCTGATACCAAGCGTTTACAGGATGCCATCGATGATTGT	291
pglA_PR8X	GCCAATCCACGGGCTCCCAGCCTGATACCAAGCGTTTACAGGATGCCATTGATGATTGT	291
pglA_Hib4	GCCAATCCACGGGCTCCCAGCCTGATACCAAGCGTTTACAGGATGCCATTGATGATTGT	300
pglA_Tem	GCCAATCCAAAGTTTCCAAGCCTGATACCAAGCGTTTACAGGATGCCATTGATGATTGT	264

pglA_9a5c	CCCGCCGGTAGCGCGGTCAAACCTAGTGATGGATAGTCATCGCAAATCCGGCTTTTGTAGT	351
pglA_PR8X	CCCGCCGGTAGCGCGGTCAAACCTAGTGATGGATAGTCATCGCAAATCCGGCTTTTGTAGC	351
pglA_Hib4	CCCGCCGGTAGCGCGGTCAAACCTAGTGATGGACAGTCATCGCAAATCCGGCTTTTGTAGT	360
pglA_Tem	CCCGCCGGTAGCGCGGTCAAACCTAGTGATCGACAGTCATCGCAAATCCGGCTTTTGTAGT	324

pglA_9a5c	GGGCCGTTGCACCTTGAAATCTGGGGTGACGTTGTGGATTGACGATGGCGTCACGCTGTTT	411
pglA_PR8X	GGGCCGTTGCACCTTGAAATCTGGGGTGACGTTGTGGATTGACGATGGCGTCACGCTGTTT	411
pglA_Hib4	GGGCCGTTGCACCTTGAAATCTGGGGTGACGTTGTGGATTGACGATGGCGTCACGCTGTTT	420
pglA_Tem	GGGCCGTTGCACCTTGAAATCTGGGGTGACGTTGTGGATTGACGAGGGTGTACGCTGTTT	384

pglA_9a5c	GCATCGCGTAACCCCAAGGACTATGACAAAGGCAATGGTACGTGCGGTACTGCAACGTCA	471
pglA_PR8X	GCATCGCGTAACCCCAAGGATTACGACAAAGGCAATGGTACGTGCGGTACTGCAACGTCA	471
pglA_Hib4	GCATCGCGTAACCCCAAGGATTACGACAAAGGCAATGGTACGTGCGGTACTGCGGAAGCTCA	480
pglA_Tem	GCATCGCGTAACCCCAAGGATTACGACAAAGGCAATGGTACGTGCGGTACTGCAACGTCA	444

	▼	
pglA_9a5c	ACCTATGAGTTCAGTTGTATGCCATTAGATCTCCGCGATCAACACCACCGGTAGCGGTAT	531
pglA_PR8X	ACCTATGAGTTCAGTTGTATGCCATT-GATCTCCGCGATCAACACCACCGGTAGCGGTAT	530
pglA_Hib4	ACCTATGAGTTCAGTTGTATGCCATT-GATCTCCGCGATCAACACCACCGATAGCGGTAT	539
pglA_Tem	ACTCATGAGTTCAGTTGTATGCCATT-GATCTCCGCGATCAACACCACCGGTAGCGGTAT	503

pglA_9a5c	CGTGGGTGGGGGTGTGATTGATGGGCGTGGCGGTAGCATTTTACTGGTGGTAAGCATGC	591
pglA_PR8X	CGTGGGTGGGGGTGTGATTGATGGGCGTGGCGGTAGCATTTTACTGGGCGTAAGCATGC	590
pglA_Hib4	CGTGGGCGGGGTGTGATTGATGGGCGTGGCGGTAGCATTTTACTGGGCGTAAGCATGC	599
pglA_Tem	CGTGGGTGGGGGTGTGATTGATGGGCGTGGCGGTAGCATTTTACTGGGCGTAAGTATGC	563

pglA_9a5c	TCGGCAGCGTACCTGGTGGGATCTGGCCTATCAGAACAAGCACCATGCACTGCATCAACA	651
pglA_PR8X	TCGGCAGCGTACCTGGTGGGATCTGGCCTATCAGAACAAGCACCATGCACTGCATCAACA	650
pglA_Hib4	TCGGCAGCGTACCTGGTGGGATCTGGCCTATCAGAACAAGCACCATGCACTGCATCAACA	659
pglA_Tem	TCGGCAGCGTACCTGGTGGGACCTGGCCTATCAGAACAAGCGCCATGCACTGCATCAACA	623

pglA_9a5c	GGTACCACGGTTGATCCAAATCAGGGGCGGTAACGACTTCACGTTGTACCGTGTGCGCCAT	711
pglA_PR8X	GGTACCACGGTTGATCCAAATCAGGGGCGGTAACGACTTCACGTTGTACCGTGTGCGCCAT	710
pglA_Hib4	GGTACCACGGTTGATCCAAATCAGGGGCGGTAACGACTTCACGTTGTACCGTGTGCGCCAT	719
pglA_Tem	GGTACCACGGTTGATCCAAATCAGGGGCGGTAACGACTTCACGTTGTACCGTGTGCGCCAT	683

pglA_9a5c	CGAAAAATGCACCTAACTTCCATGTGGTTGCTGATACGGTCAGTGGTGTACCCGCATGGGG	771
pglA_PR8X	CGAAAAATGCACCTAACTTCCATGTGGTTGCTGATACGGTCAGTGGTGTACCCGCATGGGG	770
pglA_Hib4	CGAAAAATGCACCTAACTTCCATGTGGTTGCTGATACGGTCAGTGGTGTACCCGCATGGGG	779
pglA_Tem	CGAAAAATGCACCTAACTTCCATGTGGTTGCTGATACGGTTAGCGGTGTACCCGCATGGGG	743

pglA_9a5c	CATCCGTATCCTGACCCCAAGCTTGGTCTATACCATGCCGGGGTACCCTGTCCGCCCGG	831
pglA_PR8X	CATCCGTATCCTGACCCCAAGCTTGGTCTATACCATGCCGGGGTACCCTGTCCGCCCGG	830
pglA_Hib4	CATCCGTATCCTGACCCCAAGCTTGGTCTATACCATGCCGGGGTACCCTGTCCGCCCGG	839
pglA_Tem	CATCCGCATCCTGACCCCAAGCTTGGTCTATACCATGCCGGGGTACCCTGTCCGCCCGG	803

pglA_9a5c	TACCACGCCGGATGTTGTGACGCCGGCGACATGCTTTACCCCGGACACGGTCAAGAACAC	891
pglA_PR8X	TACCACGCCGGATGTTGTGACGCCGGCGACATGCTTTACCCCGGACACGGTCAAGAACAC	890
pglA_Hib4	TACCACGCCGGATGTTGTGACGCCGGCGACATGCTTTACCCCGGACACGGTCAAGAACAC	899
pglA_Tem	TACCACGCCGGATGTTGTGACGCCGGCGACATGCTTTACCCCGGACACGGTCAAGAACAC	863

pglA_9a5c	CGATGGCTTTGATCCTGGGCAATCAAACCACGTGCTGCTAGCGTACTCCTACATCAGCAC	951
pglA_PR8X	CGATGGCTTTGATCCTGGGCAATCAAACCACGTGCTGCTAGCGTACTCCTACATCAGCAC	950
pglA_Hib4	CGATGGCTTTGATCCTGGGCAATCAAACCACGTGCTGCTAGCGTACTCCTACATCAGCAC	959
pglA_Tem	CGATGGCTTTGATCCTGGGCAATCAAACCACGTGCTGCTAGCGTACTCCTACATCAGCAC	923

pglA_9a5c	TGGGGATGACCACGTTGCCATCAAAGCGCGGGGCAAGATGCCAAGCGACGCTCTGAGCTT	1011
pglA_PR8X	TGGGGATGACCACGTTGCCATCAAAGCGCGGGGCAAGATGCCAAGCGACGCTCTGAGCTT	1010
pglA_Hib4	TGGGGATGACCACGTTGCCATCAAAGCGCGGGGCAAGATGCCAAGCGACGCTCTGAGCTT	1019
pglA_Tem	TGGGGATGACCACGTTGCCATCAAAGCGCGGGGCAAGATGCCAAGCTACGCTCTGAGCTT	983

pglA_9a5c	CTTACACAACCACTTTGGCTATGGACATGGCATGTCCATCGGTAGCGACACCGAGTCAGG	1071
pglA_PR8X	CTTACACAACCACTTTGGCTATGGACATGGCATGTCCATCGGTAGCGACACCGAGTCAGG	1070
pglA_Hib4	CTTACACAACCACTTTGGCTATGGACATGGCATGTCCATCGGTAGCGACACCGAGTCAGG	1079
pglA_Tem	CTTACACAACCACTTTGGCTATGGACATGGCATGTCCATCGGTAGCGACACCGAGTCAGG	1043

pglA_9a5c	TGTACACGATATGGAAGTCAGTGATTTGAGTATCGACGGCTTTGACAGCCCTAACAGCAA	1131
pglA_PR8X	TGTACACGATATGGAAGTCAGTGATTTGAGTATCGACGGCTTTGACAGCCCTAACAGCAA	1130
pglA_Hib4	TGTACACGATATGGAAGTCAGTGATTTGAGTATCGACGGCTTTGACAGCCCTAACAGCAA	1139
pglA_Tem	TGTACACGATATGGAAGTCAGTGATTTGAGTATCGACGGCTTTGACAGCCCTAACAGCAA	1103

pglA_9a5c	TGGCTTGCATATGAAATCTGATGCCGATCATGGCGGAGTGGTCGATCATGTCACCTACTC	1191
pglA_PR8X	TGGCTTGCATATGAAATCTGATACCGATCATGGCGGAGTGGTCGACCATGTCACCTACTC	1190
pglA_Hib4	TGGCTTGCATATGAAATCTGATACCGATCATGGCGGAGTGGTCGACCATGTCACCTACTC	1199
pglA_Tem	TGGTTTGCAGATGAAATCTGATGCCGATCATGGCGGAGTGGTCGACCATGTCACCTACTC	1163
*** *****		
pglA_9a5c	CAAAATATGCATGCGTCGCTGAAACGTCGCTGGCGTTTGACACCTTCTACAAACCTTC	1251
pglA_PR8X	CAAAATATGCATGCGTCGCTGAAACGTCGCTGGCGTTTGACACCTTCTACAAACCTTC	1250
pglA_Hib4	CAAAATATGCATGCGTCGCTGAAACGTCGCTGGCGTTTGACACCTTCTACAAACCTTC	1259
pglA_Tem	CAAAATATGCATGCGTCGCTGAAACGTCGCTGGCGTTTGACACCTTCTACAAACCTTC	1223

pglA_9a5c	CAACGGGAATTCTTATCCACTCTTCAAAAATATCGTGCTGCAAGATATTCATGTGCTTGA	1311
pglA_PR8X	CAACGGGAATTCTTATCCACTCTTCAAAAATATCGTGCTGCAAGATATTCATGTGCTTGA	1310
pglA_Hib4	CAACGGGAATTCTTATCCACTCTTCAAAAATATCGTGCTGCAAGATATTCATGTGCTTGA	1319
pglA_Tem	CAACGGGAATTCTTATCCACTCTTCAAAAATATCGTGCTGCAAGATATTCATGTGCTTGA	1283

pglA_9a5c	AAGTCTGTCTTTGGTGCTGGTCAGTGCTTTTCATGGGGATCCTCGGCAGTGGGAATAA	1371
pglA_PR8X	AAGTCTGTCTTTGGTGCTGGTCAGTGCTTTTCATGGGGATCCTCGGCAGTGGGAATAA	1370
pglA_Hib4	AAGTCTGTCTTTGGTGCTGGTCAGTGCTTTTCATGGGGATCCTCGGCAGTGGGAATAA	1379
pglA_Tem	AAGTCTGTCTTTGGTGCTGGTCAGTGCTTTTCATGGGGATCCTCGGCAGTGGGAATAA	1343

pglA_9a5c	TCTGCCAATGACGCTGAGTCTGGATAACGTCGTATTCGACGGTTTTTTGCCAACATTGAT	1431
pglA_PR8X	TCTGCCAATGACGCTGAGTCTGGATAACGTCGTATTCGACGGTTTTTTGCCAACATTGAT	1430
pglA_Hib4	TCTGCCAATGACGCTGAGTCTGGATAACGTCGTATTCGACGGTTTTTTGCCAACATTGAT	1439
pglA_Tem	TCTGCCAATGACGCTGAGCATGGATAACGTCGTATTCGACGGTTTTTTGCCAACATTGAT	1403

pglA_9a5c AGCGCCTCCCAGCAGTGTCTGATTTCGCCAATCCGCAGGCAGTGCATTTTCACTTCGGTCC 1491
pglA_PR8X AGCGCCTCCCAGCAGTGTCTGATTTCGCCAATCCGCAGGCAGTGCATTTTCACTTCGGTCC 1490
pglA_Hib4 AGCGCCTCCCAGCAGTGTCTGATTTCGCCAATCCGCAGGCAGTGCATTTTCACTTCGGTCC 1499
pglA_Tem AGCGCCTCCCAGCAGTGTCTGATTTCGCCAATCCGCAGGCAGTGCATTTTCACTTCGGTCC 1463

pglA_9a5c CGGTCCGGTCAGTTTCGCTCCTTTGATTACCCCTCGGTTGCTTACGATGTGACTGTGAG 1551
pglA_PR8X CGGTCCGGTCAGTTTCGCTCCTTTGATTACCCCTCGGTTGCTTACGATGTGACTGTGAG 1550
pglA_Hib4 CGGTCCGGTCAGTTTCGCTCCTTTGATTACCCCTCGGTTGCTTACGATGTGACTGTGAG 1559
pglA_Tem CGGTCCGGTCAGTTTCGCTCCTTTGATTACCCCTCGGTTGCTTACGATGTGACTGTGAG 1523

pglA_9a5c TGGAAGTCCAGGTGTGGGCAACCCATACGATTGTTTCGGCGGCATTATCAATTTTCCAG 1611
pglA_PR8X TGGAAGTCCAGGTGTGGGCAACCCATACGATTGTTTCGGCGGCATTATCAATTTTCCAG 1610
pglA_Hib4 TGGAAGTCCAGGTGTGGGCAACCCATACGATTGTTTCGGCGGCATTATCAATTTTCCAG 1619
pglA_Tem TGGAAGTCCAGGTGTGGGCAACCCATACGATTGTTTCGGCGGCATTATCAATTTTCCAG 1583

pglA_9a5c CGTATTTCTCTGATTGCGCTATCTAA 1636
pglA_PR8X CGTATTTCTCTGATTGCGCTATCTAA 1635
pglA_Hib4 CGTATTTCTCTGATTGCGCTATCTAA 1644
pglA_Tem CGTATTTCTCTGATTGCGCTATCTAA 1608

B)

pglA_9a5c MNLDRFLPLVFLGLHCFAMGG--TADVFRGSQPSLSDLSEPVSVKTLWGEVQRPSLPTH 57
pglA_PR8X MNLDRFLPLVFLGLHCFAMGG--TADVFRGSQPSLSDLSEPVSVKTLWGEVQRPSLPTH 57
pglA_Hib4 MNLDRFLPLVFLGLHCFAMGGTADTADVFRVSQPSLSDLSEPVSVKTLWGEVQRPSLPTH 60
pglA_Tem -----MFGLHCFAMGG--TADVFRVSQPSLSDLSEPVSVKTLWGEVQRPSLPTH 48
:***** * * *

pglA_9a5c VCTVLPARLTPEHGSLDALDANPRASKPDTKRLQDAIDDCPAGSAVKLVMDSHRKSGLS 117
pglA_PR8X VCTVLPARLTPEHGSLDALDANPRASQPDTKRLQDAIDDCPAGSAVKLVMDSHRKSGLS 117
pglA_Hib4 VCTVLPARLTPEHGSLDALDANPRASQPDTKRLQDAIDDCPAGSAVKLVMDSHRKSGLS 120
pglA_Tem VCTVLPARLTPKHGSIDALDANPKVSKPDTKRLQDAIDDCPAGSAVKLVMDSHRKSGLS 108
*****:*. :*****:*. :*****:*****:*****

pglA_9a5c GPLHLKSGVTLWIDDGVTLFASRNPKDYDKNGTCGTATSTYEFSCMPLDLRDQHHR-- 174
pglA_PR8X GPLHLKSGVTLWIDDGVTLFASRNPKDYDKNGTCGTATSTYEFSCMPLISAINTTSGSI 177
pglA_Hib4 GPLHLKSGVTLWIDDGVTLFASRNPKDYDKNGTCGTATSTYEFSCMPLISAINTTSGSI 180
pglA_Tem GPLHLKSGVTLWIDEGVTLFASRNPKDYDKNGTCGTATSTYEFSCMPLISAINTTSGSI 168
*****:*****:***:*****:***

pglA_PR8X VGGGVIDGRGGSILTGKHKARQRTWWDLAYQNKHHALHQQVPRLIQIRGGNDFTLYRVAI 237
pglA_Hib4 VGGGVIDGRGGSILTGKHKARQRTWWDLAYQNKHHALHQQVPRLIQIRGGNDFTLYRVAI 240
pglA_Tem VGGGVIDGRGGSILTGKHYARQRTWWDLAYQNKHHALHQQVPRLIQIRGGNDFTLYRVAI 228
*****:*****:*****:*****:*****

pglA_PR8X ENAPNFHVVDTVSGVTAWGIRILTPSLVYTTPGYHCPPGTTDPDVVTPATCFTPDVTKNT 297
pglA_Hib4 ENAPNFHVVDTVSGVTAWGIRILTPSLVYTTPGYHCPPGTTDPDVVTPATCFTPDVTKNT 300
pglA_Tem ENAPNFHVVDTVSGVTAWGIRILTPSLVYTTPGYHCPPGTTDPDVVTPATCFTPDVTKNT 288

pglA_PR8X DGFDPGQSNHVLLAYSISTGDDHVAIKARGKMPDALSFLHNYFGYGHGMSIGSDTESG 357
pglA_Hib4 DGFDPGQSNHVLLAYSISTGDDHVAIKARGKMPDALSFLHNYFGYGHGMSIGSDTESG 360
pglA_Tem DGFDPGQSNHVLLAYSISTGDDHVAIKARGKMPDALSFLHNYFGYGHGMSIGSDTESG 348
*****:*****:*****:*****

pglA_PR8X VHDMEVSDLSIDGFDPSPNSNGLHMKSDTDHGGVVDHVTYSKICMRRLKRPLAFDTFYKPS 417
pglA_Hib4 VHDMEVSDLSIDGFDPSPNSNGLHMKSDTDHGGVVDHVTYSKICMRRLKRPLAFDTFYKPS 420
pglA_Tem VHDMEVSDLSIDGFDPSPNSNGLQMKSDADHGGVVDHVTYSKICMRRLKRPLAFDTFYKPS 408
*****:***:*****:*****

pglA_PR8X NGNSYPLFKNIVLQDIHVLESVPVFGAGQLLFMGILGSGNNLPMTLSDNVVFDGFLPTLI 477
pglA_Hib4 NGNSYPLFKNIVLQDIHVLESVPVFGAGQLLFMGILGSGNNLPMTLSDNVVFDGFLPTLI 480
pglA_Tem NGNSYPLFKNIVLQDIHVLESVPVFGAGQLLFMGILGSGNNLPMTLSDNVVFDGFLPTLI 468
*****:*****

pglA_PR8X APPSSVVFANPQAVHFHFGPGPVSFAPLITPSVAYDVTVSGSPGVGNPYDCSAAFIDFSS 537
pglA_Hib4 APPSSVVFANPQAVHFHFGPGPVSFAPLITPSVAYDVTVSGSPGVGNPYDCSAAFIDFSS 540
pglA_Tem APPSSVVFANPQAVHFHFGPGPVSFAPLITPSVAYDVTVSGSPGVGNPYDCSAAFIDFSS 528
*****:***

pglA_PR8X VFPD SPI 544
pglA_Hib4 VFPD SPI 547
pglA_Tem VFPD SPI 535

Figura S2. Esquema da organização de genes dos plasmídeos pXF39 da cepa Pr8x e pXF51 das cepas 9a5c e J1a12. O esquema foi preparado a partir da anotação disponível no IMG-ER. Os números no topo da figura indicam algumas das CDS de pXF39 listadas na Tabela 16. O retângulo tracejado em vermelho indica, aproximadamente, a região de pXF51 ausente em pXF39.

Figura S3. Alinhamento da sequência de nucleotídeos de XF1952 das cepas 9a5c, J1a12, Pr8x e Hib4. Está mostrada apenas parte do alinhamento para destacar a inserção de 21nt em fase na sequência desta CDS em Hib4.

```

XF1952_Hib      GCAGAAACACGTCGGACGTTAGATGTTCTGGCGCATGAGGCGATTAGCAATTTTACTGCA 1020
XF1952_Pr      GCAGAAACACGTCGGACGTTAGATGTTCTGGCGCATGAGGCGATTAGCAATTTTACTGCA 909
XF1952_9a      GCAGAAACACGTCGGACGTTAGATGTTCTGGCGCATGAGGCGATTAGCAATTTTACTGCA 1020
XF1952_J1      GCAGAAACACGTCGGACGTTAGATGTTCTGGCGCATGAGGCGATTAGCAATTTTACTGCA 1020
*****

XF1952_Hib      GCCCGTGAATGCTTTATTGCTTTTCATTGAGAGTAATTGGGAGCAAGCGCGCTTGCTTAAG 1080
XF1952_Pr      GCCCGTGAATGCTTTATTGCTTTTCATTGAGAGTAATTGGGAGCAAGCGCGCTTGCTTAAG 969
XF1952_9a      GCCCGTGAATGCTTTATTGCTTTTCATTGAGAGTAATTGGGAGCAAGCGCGCTTGCTTAAG 1080
XF1952_J1      GCCCGTGAATGCTTTATTGCTTTTCATTGAGAGTAATTGGGAGCAAGCGCGCTTGCTTAAG 1080
*****

XF1952_Hib      GTACCAGGTTTGCTCAGCGAAGTTGGCGGTGCGTTGTCAATTCTTGAGTTGTACACAGCT 1140
XF1952_Pr      GTACCAGGTTTGCTCAGCGAAGTTGGCGGTGCGTTGTCAATTCTTGAGTTGTACACAGCT 1029
XF1952_9a      GTACCAGGTTTGCTCAGCGAAGTTGGCGGTGCGTTGTCAATTCTTGAGTTGTACACAGCT 1140
XF1952_J1      GTACCAGGTTTGCTCAGCGAAGTTGGCGGTGCGTTGTCAATTCTTGAGTTGTACACAGCT 1140
*****

XF1952_Hib      GCAAGCTATCTGGAGGGGTGTACGCCGTTATATCGAGTGTGAGCTAATTACCAACGTCGT 1200
XF1952_Pr      GCAAGCTATCTGGAGGGGTGTACGCCGTTATATCGAGTGTGAGCTAATTACCAACGTCGT 1089
XF1952_9a      GCAAGCTATCTGGAGGGGTGTACGCCGTTATATCGAGTGTGAGCTAATTACCAACGTCGT 1200
XF1952_J1      GCAAGCTATCTGGAGGGGTGTACGCCGTTATATCGAGTGTGAGCTAATTACCAACGTCGT 1200
*****

XF1952_Hib      GTGCTGCTGCTCGTCAGCTGGACACTCTGGCTGATGCGATGGCCAGTTTGAATACTAC 1260
XF1952_Pr      GTGCTGCTGCTCGTCAGTTGGACACTCTGGCTGATGCGATGGCCAGTTTGAATACTAT 1149
XF1952_9a      GTGCTGCTGCTCGTCAGTTGGACACTCTGGCTGATGCGATGGCCAGTTTGAATACTAC 1260
XF1952_J1      GTGCTGCTGCTCGTCAGTTGGACACTCTGGCTGATGCGATGGCCAGTTTGAATACTAT 1260
*****

XF1952_Hib      TTAGAAGCGTTGCGTGAGAGACGTCCTGGACGTGAGGACATACTCGAAATCACACGTAAC 1320
XF1952_Pr      TTAGAAGCGTTGCGTGAGAGACGTCCTGGACGTGAGGACATACTCGAAATCACACGTAAC 1209
XF1952_9a      TTAGAAGCGTTGCGTGAGAGACGTCCTGGACGTGAGGACATACTCGAAATCACACGTAAC 1320
XF1952_J1      TTAGAAGCGTTGCGTGAGAGACGTCCTGGACGTGAGGACATACTCGAAATCACACGTAAC 1320
*****

XF1952_Hib      AGTCTGGAAGCACTCTCTTACTGGCCCATCCCTGCTGAAGAGGGGGATAGCGTTGGTTCC 1380
XF1952_Pr      AGTCTGGAAGCACTGTCTTACTGGCCCATCCCTGCTGAAGAGGGGGATAGCGTTGGTTCC 1269
XF1952_9a      AGTCTGGAAGCACTGTCTTACTGGCCCATCCCTGCTGAAGAGGGGGATAGCGTTGGTTCC 1380
XF1952_J1      AGTCTGGAAGCACTGTCTTACTGGCCCATCCCTGCTGAAGAGGGGGATAGCGTTGGTTCC 1380
*****

XF1952_Hib      CAACAGAGTACAGATGTGCTGACTTTCTCTGCTGTGGCTACAGAATCTATGTACACCTGAC 1440
XF1952_Pr      CAACAGAGTACAGATGTGCTGACTTTCTCTGCT-----CCTGAC 1308
XF1952_9a      CAACAGAGTACAGATGTGCTGACTTTCTCTGCT-----CCTGAC 1419
XF1952_J1      CAACAGAGTACAGATGTGCTGACTTTCTCTGCT-----CCTGAC 1419
*****

XF1952_Hib      GAAGCGCTACCAGAACTGACTGTGGTTGAAGATGGAACGGTTACCTTTTCAGGAGGCGGTA 1500
XF1952_Pr      GAAGCGCTACCAGAACTGACTGTGGTTGAAGATGGAACGGTTACCTTTTCAGGAGGCGGTA 1368
XF1952_9a      GAAGCGCTACCAGAACTGACTGTGGTTGAAGATGGAACGGTTACCTTTTCAGGAGGCGGTA 1479
XF1952_J1      GAAGCGCTACCAGAACTGACTGTGGTTGAAGATGGAACGGTTACCTTTTCAGGAGGCGGTA 1479
*****

XF1952_Hib      TCACAATCCCTTACATTAGTTGAGCCAGAAAGCGCCGACATCTTCGACGCATCGTTGAGG 1560
XF1952_Pr      TCACAATCCCTTACATTAGTTGAGCCAGAAAGCGCCGACATCTTCGACGCATCGTTGAGG 1428
XF1952_9a      TCACAATCCCTTACATTAGTTGAGCCAGAAAGCGCCGACATCTTCGACGCATCGTTGAGG 1539
XF1952_J1      TCACAATCCCTTACATTAGTTGAGCCAGAAAGCGCCGACATCTTCGACGCATCGTTGAGG 1539
*****

XF1952_Hib      GAGCTGGTGCAGCCCATTTGTTACTGATCTATATGAAGGAGATCGACGCAGCGAGTTGCCC 1620
XF1952_Pr      GAGCTGGTGCAGCCCATTTGTTACTGATCTATATGAAGGAGATCGACGCAGCGAGTTGCCC 1488
XF1952_9a      GAGCTGGTGCAGCCCATTTGTTACTGATCTATATGAAGGAGATCGACGCAGCGAGTTGCCC 1599
XF1952_J1      GAGCTGGTGCAGCCCATTTGTTACTGATCTATATGAAGGAGATCGACGCAGCGAGTTGCCC 1599
*****

XF1952_Hib      ACTGAATCCATGCCTGACGATGAGCCGGTTATCCAAAGTATACCTACTGTTGTTGCACCG 1680
XF1952_Pr      ACTGAATCCATGCCTGACGATGAGCCGGTTATCCAAAGTATACCTACTGTTGTTGCACCG 1548
XF1952_9a      ACTGAATCCATGCCTGACGATGAGCCGGTTATCCAAAGTATACCTACTGTTGTTGCACCG 1659
XF1952_J1      ACTGAATCCATGCCTGACGATGAGCCGGTTATCCAAAGTATACCTACTGTTGTTGCACCG 1659
*****

```

8. Súmula Curricular

Paulo Marques Pierry

Nascimento: 18/05/1987 – Santos / SP - Brasil

Formação Acadêmica

- 2005 - 2008** Graduação em Ciências Biológicas
Licenciatura
Bacharelado com ênfase em Biotecnologia
Bacharelado com ênfase em Meio Ambiente
Universidade Católica de Santos, UNISANTOS, Santos, São Paulo, Brasil
- 2009 -2012** Curso de Mestrado, Programa de Ciências Biológicas (Bioquímica).
Universidade de São Paulo, USP, São Paulo, Brasil
Orientadora: Profa. Dra. Aline Maria da Silva
Bolsista da Fundação de Amparo à Pesquisa do Estado de São Paulo
-

Formação complementar

- 2006-2008** Iniciação Científica
Universidade Católica de Santos, UNISANTOS, Santos, São Paulo, Brasil
Orientadoras: Profa. Dra. Adriana Florentino de Souza
Profa. Dra. Kátia Maria Gomes Machado
Profa. Dra. Rosângela Ballego Campanã
Bolsista do Programa de Iniciação Científica (PROIN) – Universidade Católica de Santos
-

Prêmios

2008 Prêmio Bolsa Mérito pelo ótimo desempenho nas disciplinas do curso de Licenciatura e Bacharelado do curso de Ciências Biológicas da UNISANTOS.

Comunicações em Congressos

Pierry, PM; da Silva, PIP; Santana, WO; Beckedorff, FCF; Amaral, MS; Teixeira, DC; Almeida, LGP; Vasconcelos, ATR; Verjovski-Almeida, S.; da Silva, AM. Sequenciamento e comparação do genoma de uma cepa não-virulenta de *Xylella fastidiosa*. In: 56º Congresso Brasileiro de Genética, 2010, Guarujá - SP.

Pierry, PM; Campanha, RB; Machado, KMG; Souza, AF. Partial Purification and Characterization of Laccases from *Psilocybe castanella*. In: XXXVIII Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular, 2009, Águas de Lindóia, SP.

Pierry, PM; Santi-Júnior, C; Souza, AF; Machado, KMG. Degradation of textile effluents by laccases. In: XXXVII Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular - SBBq e XI Congresso da PABMB, 2008, Águas de Lindóia, SP.

Pierry, PM; Souza, AF; Machado, KMG. Produção de atividade de lacase por *Psilocybe castanella* e avaliação de seu potencial de degradação de corantes têxteis reativos. In: 5º Congresso Brasileiro de Micologia, 2007, Recife.

Santi-Júnior, C; Pierry, PM; Souza, AF; Machado, KMG. Degradação de corantes têxteis reativos por lacases de fungos basidiomicetos. In: IV EPOA, Encontro sobre Aplicações Ambientais de Processos Oxidativos Avançados, 2007, Cubatão.