UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE QUÍMICA

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS
(BIOQUÍMICA)

PATRÍCIA ISABELA PESSOA DA SILVA

**Identificação de genes com expressão modulada por
estreptomicina e de genes associados à virulência e
patogenicidade em *Xylella fastidiosa*.**

São Paulo

Data de depósito SPG: 12/02/2010

PATRÍCIA ISABELA PESSOA DA SILVA

**Identificação de genes com expressão modulada por
estreptomicina e de genes associados à virulência e
patogenicidade em *Xylella fastidiosa***

Tese apresentada ao Instituto de
Química da Universidade de São Paulo
para a obtenção do Título de Doutor em
Ciências, área Bioquímica.

Orientadora: Profa. Dra. Aline Maria da Silva

São Paulo

2010

AUTORIZO A REPRODUÇÃO E A DIVULGAÇÃO DE TODO OU PARTE DESTE 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

da Silva, Patrícia Isabela Pessoa

Identificação de genes com expressão modulada por estreptomicina e de genes associados à virulência e patogenicidade em *Xylella fastidiosa* / Patrícia Isabela Pessoa da Silva. --São Paulo, 2010.

Orientadora: Aline Maria da Silva.

138 f.

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

Descritores: 1. Fitopatígeno. 2. Clorose Variegada dos Citros. 3. Estreptomicina. 4. Perfil de Expressão Gênica. 5. Pirosequenciamento. 6. Genômica comparativa. I. da Silva, Aline Maria II. Universidade de São Paulo. Instituto de Química. Programa de Pós-Graduação em Ciências Biológicas (Bioquímica). III. Título.

UNIVERSIDADE DE SÃO PAULO

“Identificação de genes com expressão modulada por estreptomicina e de genes associados à virulência e patogenicidade em *Xylella fastidiosa*”

Patrícia Isabela Pessoa da Silva

Tese de Doutorado submetida ao Instituto de Química da Universidade de São Paulo como parte dos requisitos necessários à obtenção do grau de Doutor em Ciências - Área: Bioquímica.

Aprovado(a) por:

Profa. Dra. Aline Maria da Silva
(Orientadora e Presidente)

Profa. Dra. Suely Lopes Gomes
IQ - USP

Profa. Dra. Regina Lúcia Baldini
IQ - USP

Prof. Dr. Julio Cezar Franco de Oliveira
UNIFESP - Diadema

Profa. Dra. Alessandra Alves de Souza
IAC

SÃO PAULO
23 DE ABRIL DE 2010

A meus pais, irmãs e Henrique.

Agradecimentos

Agradeço primeiramente a Deus, pela força e por guiar meus caminhos. Aos meus pais, Dalva e Valdo, por todos os ensinamentos, valores e dedicação. Às minhas irmãs, Renata e Fernanda, por serem presentes em minha vida e por serem pessoas com quem verdadeiramente sempre poderei contar. Ao Henrique, por existir e tornar nossas vidas mais felizes. Ao Chico, por ter facilitado em todos os sentidos essa minha longa jornada. Aos tios (Darcy, Ademar, Déa, Aparecida, Piedade, Cleiner e Rose) e primos (Kátia, Daniela, Anna Clara e Bruno), pelos bons momentos que me proporcionaram cada vez que retornei à minha cidade, fazendo com que eu me sentisse importante e amada. A Cida e Nicolau, pelo aconchego e refúgio que me proporcionaram nesses anos de vida paulistana.

À professora Aline Maria da Silva pela orientação, ensinamentos e enorme compreensão nos momentos difíceis. Sua atuação ativa e empenho na conclusão desse trabalho foram, para mim, a melhor demonstração de seu interesse e cuidado como orientadora.

A Thaís e Erica, que fizeram com que eu soubesse o verdadeiro significado da palavra amizade. A Nathália, por tornar a hora do almoço um evento imperdível! Aos amigos: Paulo Paiva, Diorge, Fabiane, Daniela Beton, Daniela Gonzalez, Ciça, Xanda, Érica Bandeira (a mãe de Kril!) e Luciana; pelos momentos divertidos que fizeram essa caminhada mais leve e agradável. Aos integrantes dos laboratórios dos professores Walter Terra, Suely L. Gomes, Gláucia Souza e Sergio Verjovski-Almeida, que sempre me receberam com um sorriso e disposição para ajudar em tudo que precisei.

Aos técnicos do meu laboratório, Alexandre Sanchez e Ivone Pereira, que tanto facilitam o bom andamento de nossos experimentos, além de terem se tornado grandes amigos. Aos colegas do laboratório, Andréa Fogaça, Ana Carolina Simões, Daniela Gonzalez, Fernanda Almeida, Juliana Moreira, Layla Martins, Leandro Moreira, Maurício Barlera, Paulo Pierry, Paulo Zaini, Renato Raposo e Wesley Santana pela convivência, participação, ensinamentos e por sempre estarem dispostos a me ajudar.

A Adriana Matsukuma e Denise Yamamoto por toda a ajuda e dedicação no laboratório de microarranjos do IQ-USP.

Ao Felipe Beckedorff e Murilo Sena por toda grande e fundamental colaboração no sequenciamento do genoma da J1a12 realizado no Centro Avançado de Tecnologias em Genômica (CATG) do IQ-USP.

Ao professor Sergio Verjovski-Almeida pelo apoio e orientações cruciais para o pirosequenciamento realizado no CATG e por possibilitar acesso irrestrito à infra-estrutura de seu laboratório durante o desenvolvimento do meu doutorado.

Ao Prof. Henrique Ferreira (UNESP, São José do Rio Preto) pelos ensinamentos na nossa tentativa de implantar o sistema de duplo-híbrido em bactéria. A Dra. Diva Teixeira (Fundecitrus) pela ajuda nos experimentos de infecção de tabaco. Ao Dr. Fabio Nakano (Instituto Butantan) pela colaboração nas análises integradas da expressão gênica.

Ao Luiz Gonzaga Paula de Almeida e à Dra. Ana Tereza Ribeiro Vasconcelos do Laboratório de Bioinformática do LNCC pela valiosa colaboração na montagem dos genomas de *X. fastidiosa*.

Ao professor Bayardo Torres, pelo exemplo do que deve ser um verdadeiro Mestre e pela oportunidade de aprender a arte de ensinar.

A todos aqueles que de alguma forma contribuíram para o meu crescimento e aprendizado nessa longa estrada de vida acadêmica.

A toda comunidade do IQ-USP por construir um espaço agradável para se trabalhar.

À FAPESP e ao CNPq pelo apoio financeiro.

Resumo

DA SILVA, P.I.P. **Identificação de genes com expressão modulada por estreptomicina e de genes associados à virulência e patogenicidade em *Xylella fastidiosa***. 2010. 138 f. Tese de Doutorado - Instituto de Química, Universidade de São Paulo, São Paulo.

Em concentrações subletais, agentes antimicrobianos modulam a expressão gênica bacteriana, sendo que o conjunto de genes que é modulado depende tanto da cepa bacteriana, como da natureza do agente antimicrobiano. Neste trabalho, avaliamos o perfil de expressão gênica de *Xylella fastidiosa* cepa 9a5c em resposta ao tratamento por até 60 minutos com dose subletal do antibiótico estreptomicina. Esta é uma cepa virulenta, originalmente isolada de laranjeiras com sintomas de clorose variegada dos citros. A hibridização de microarranjos de DNA representando 2608 das 2848 sequências codificadoras (CDS) previamente anotadas no genoma desta cepa, revelou que 136 CDS apresentaram expressão gênica diferencial em resposta à exposição à estreptomicina, sendo que destas 109 foram negativamente moduladas e 27 positivamente moduladas. Realizamos, também, ensaios de PCR quantitativo precedido de transcrição reversa (RT-qPCR) de 21 CDS para confirmar a modulação observada na análise global da expressão gênica. O perfil de expressão gênica de *X. fastidiosa* em resposta à estreptomicina foi analisado de forma integrada com outros perfis de expressão gênica desta bactéria. Entre as CDS positivamente moduladas, destacamos aquelas codificadoras das chaperoninas GroEL e GroES, que estão associadas a resposta de choque térmico, e CDS associadas à tradução, tais como proteínas ribossomais e fatores de tradução. Interessantemente, a exposição à estreptomicina induz a expressão da CDS que codifica poligalacturonase, que é um fator de virulência em algumas cepas de *X. fastidiosa*. Por outro lado, o tratamento com estreptomicina promoveu a modulação negativa de CDS relacionadas à formação e manutenção de biofilme ao contrário do observado quando estas bactérias foram submetidas ao tratamento com gomesina, um peptídeo antimicrobiano. O conjunto destas observações sugere que a exposição à dose subletal de estreptomicina possa promover um fenótipo de maior virulência, contrariamente ao efeito previamente observado com a gomesina. Neste trabalho, também descrevemos o pirosequenciamento do genoma da cepa J1a12 de *Xylella fastidiosa*, que exibe fenótipo menos virulento em citros e tabaco em relação à cepa 9a5c. A comparação da sequência genômica destas duas cepas confirma diferenças anteriormente observadas utilizando-se microarranjos de DNA e destaca genes potencialmente importantes para virulência de *Xylella fastidiosa*.

Palavras chave: Fitopatógeno, Clorose Variegada dos Citros, Estreptomicina, Perfil de expressão gênica, Pirosequenciamento, Genômica comparativa.

Abstract

DA SILVA, P.I.P. **Identification of genes modulated by streptomycin and of genes related to virulence and pathogenicity in *Xylella fastidiosa***. 2010. 138 f. PhD Thesis - Instituto de Química, Universidade de São Paulo, São Paulo.

At sublethal concentrations, antimicrobials compounds modulate bacterial gene expression and the gene set that is modulated depends not only on the bacterial strain but also on the nature of antimicrobial agent. In this study, we evaluated changes in gene expression profile of *Xylella fastidiosa* strain 9a5c exposed up to 60 min to sublethal concentration of streptomycin. This a virulent strain originally isolated from orange trees with symptoms of citrus variegated chlorosis. Hybridization of DNA microarrays representing 2,608 out of 2848 coding sequences (CDS) previously annotated in strain 9a5c genome revealed 136 CDS differentially expressed upon streptomycin treatment. Of which 109 were down-regulated and 27 up-regulated. Differential expression for a subset of 21 CDS was further evaluated by reverse transcription–quantitative PCR (RT-qPCR). In addition, we performed an integrated analysis of the gene expression profile of *X. fastidiosa* in response to streptomycin along with other gene expression profiles available for this bacterium. Among the up-regulated CDS, we highlight those encoding chaperonins GroEL and GroES, which are associated with heat shock response, and those CDS related to translation, such as ribosomal proteins and translation factors. Interestingly, exposure to streptomycin induces the expression of a CDS encoding polygalacturonase, which is a virulence factor for some *X. fastidiosa* strains. Furthermore, treatment with streptomycin down-regulates some CDS related to biofilm formation oppositely to treatment with gomesin, an antimicrobial peptide. Together, these observations suggest that exposure to sublethal dose of streptomycin might promote a higher virulent phenotype, in contrast to the effect previously observed with gomesin. In the present work, we also describe the pyrosequencing of J1a12 genome, a *X. fastidiosa* strain that exhibits a less virulent phenotype in citrus and tobacco if compared to strain 9a5c. A comparison of genome sequences of these two strains confirms differences previously observed using DNA microarrays and highlights important genes for virulence of *X. fastidiosa*.

Keywords: Phytopathogen, Citrus Variegated Chlorosis, Streptomycin, Expression Profiling, Pyrosequencing, Comparative Genomics.

Lista de Abreviaturas e Siglas

ATP: adenosina trifosfato

AMP: Peptídeo antimicrobiano

BLAST: *Basic Local Alignment Search Tool*

CATG: Centro Avançado de Tecnologias em Genômica, IQUSP

cDNA: DNA complementar

CDS: sequência codificadora

CVC: Clorose Variegada dos Citros

DEPC: dietilpirocarbonato

DMSO: dimetil sulfóxido

DNA: ácido desoxirribonucleico

Dnase: desoxirribonuclease

dNTP: desoxirribonucleotídeo 5' trifosfato

DO: densidade óptica

DSF: *Diffusible Signaling Factor*

DTT: ditionitrito

EDTA: ácido etilenodiaminotetracético

EPS: polissacarídeo extracelular

Fundecitrus: Fundo de Defesa da Citricultura

GEO: *Gene Expression Omnibus*

kb: quilobase

kDa: quilo Dalton

LNCC: Laboratório Nacional de Computação Científica

mA: miliampére

MAD: desvio absoluto da mediana

MBC: concentração bactericida mínima

mJ: milijoule

mRNA: RNA mensageiro

MTM Dens: densidade de fluorescência média removida de artefatos

NCBI: *National Center for Biotechnology Information*

pb: pares de base

PCR: reação em cadeia da polimerase

PD: Doença de Pierce

qPCR: PCR quantitativa

RAST: *Rapid Annotation using Subsystem Technology*

RNA: ácido ribonucleico

Rnase: Ribonuclease A

rpm: rotações por minuto

RT: transcrição reversa

RT-qPCR: qPCR precedida por RT

SABIA: *System for Automated Bacterial (genome) Integrated Annotation*

SDS: dodecil sulfato de sódio

SSC: tampão citrato de sódio

ssDNA: *single-stranded DNA*

TAE: Tampão Tris-Acetato contendo EDTA

TBS: Solução salina tamponada com Tris

TE: Tampão Tris contendo EDTA

T_m: temperatura de fusão

Tris: tris-(hidroximetil)-aminometano

U: unidade de atividade enzimática

UV: ultravioleta

Lista de Figuras e Tabelas

	Pág
Figura 1. Imagens de agregados de células e do biofilme formados por <i>X. fastidiosa</i> na planta e no inseto-vetor	17
Figura 2: Árvore filogenética de isolados de <i>Xylella fastidiosa</i>	18
Figura 3: Sintomas típicos da Clorose Variegada dos Citros	19
Figura 4: Sintomas típicos da Doença de Pierce em videiras	19
Figura 5. Avaliação da integridade das amostras de RNA por eletroforese em gel de agarose em condições desnaturantes	30
Figura 6. Avaliação da integridade do RNA total no Bioanalyzer (Agilent Technologies).	31
Figura 7. Infecção de <i>Nicotiana clevelandii</i> com <i>Xylella fastidiosa</i>	37
Figura 8. Análise eletroforética de amostra de DNA genômico de <i>X. fastidiosa</i>	39
Figura 9. Quantificação das amostras de DNA genômico de <i>X. fastidiosa</i> com PicoGreen (Invitrogen)	40
Figura 10. Genotipagem das cepas 9a5c e J1a12	42
Figura 11. Ensaio de viabilidade celular de <i>X. fastidiosa</i> exposta a doses subletais de estreptomicina.	46
Figura 12. Perfis de expressão de CDS selecionadas em resposta ao tratamento com dose subletal de estreptomicina avaliados por RT-qPCR	54
Figura 13. Comparação de perfis de expressão <i>Xylella fastidiosa</i> submetida a diferentes condições experimentais	68
Figura 14. Distribuição das sequências geradas no sequenciador 454 por tamanho em bases	74
Figura 15. Comparação das sequências de colicinas V da cepa 9a5c e J1a12	79
Figura 16. Resultado da busca no <i>GenBank</i> por sequências similares ao <i>contig 374</i>	79
Figura 17. Anotação preliminar do <i>contig 374</i>	80
Tabela 1. Genomas sequenciados de cepas de <i>Xylella fastidiosa</i>	20
Tabela 2. Iniciadores utilizados nas reações de RT-qPCR	35
Tabela 3. Preparo da curva padrão <i>high-range</i> para quantificação de DNA genômico com o Quan-iT PicoGreen (Invitrogen)	40
Tabela 4. Determinação do tamanho e do número de regiões da placa carregadas com as bibliotecas de DNA	41
Tabela 5. PCR para genotipagem de 9a5c e J1a12	42
Tabela 6. Perfil de suscetibilidade da <i>X. fastidiosa</i> a diferentes antibióticos	46
Tabela 7. Total de CDS de <i>X. fastidiosa</i> expressas e diferencialmente expressas em resposta ao tratamento com 1µg/mL de estreptomicina	48
Tabela 8. CDS moduladas positivamente em resposta ao tratamento com 1µg/mL de estreptomicina	49
Tabela 9. CDS moduladas negativamente em resposta ao tratamento com 1µg/mL de	50

estreptomicina

Tabela 10. Comparação dos dados de microarranjo e de RT-qPCR para 25 CDS	55
Tabela 11. Avaliação do perfil de expressão de 11 CDS utilizando RT-qPCR	57
Tabela 12. Comparação do perfil de expressão das CDS moduladas positivamente pela estreptomicina aos seus respectivos perfis quando <i>X.fastidiosa</i> é submetida a outras condições experimentais	64
Tabela 13. Condições experimentais incluídas na análise integrada de expressão gênica de <i>X. fastidiosa</i>	67
Tabela 14. Avaliação da virulência de <i>Xylella fastidiosa</i> cepas 9a5c e J1a12 inoculadas em <i>Nicotiana clevelandii</i>	71
Tabela 15. Desempenho do pirosequenciamento de DNA cepas J1a12 e 9a5c	73
Tabela 16. Montagem das sequências das cepas J1a12 e 9a5c ^b	75
Tabela 17. Comparação entre genomas de J1a12 e 9a5c: CDS previamente identificadas como altamente divergentes ou divergentes	78
Tabela S1. Comparação do perfil de expressão das CDS moduladas negativamente pela estreptomicina aos seus respectivos perfis quando <i>X. fastidiosa</i> é submetida a outras condições experimentais	99
Tabela S2. Comparação do perfil de expressão das CDS anotadas no genoma de <i>X.fastidiosa</i> 9a5c em diversas condições experimentais. (disponível no CD-ROM complementar a esta Tese).	99
Tabela S3. Comparação do perfil de expressão das CDS selecionadas em diversas condições experimentais.	99
Tabela S4. Produto predito das CDS selecionadas em diversas condições experimentais e listadas na Tabela S3.	99
Tabela S5. Resultado do mapeamento dos <i>reads</i> do sequenciamento do genoma da cepa J1a12 no genoma referência (cepa 9a5c). (disponível no CD-ROM complementar a esta Tese)	99
Figura S1. Alinhamento das sequências de nucleotídeos do <i>contig</i> 374 e do plasmídeo pVEIS01. As sequências foram alinhadas utilizando-se o programa CLUSTAL. (*) indica nucleotídeos idênticos. O alinhamento completo está disponível no CD-ROM complementar a esta Tese.	107

Sumário

1. Introdução	15
1.1. O fitopatógeno <i>Xylella fastidiosa</i>	15
1.2. Genomas de <i>Xylella fastidiosa</i>	20
1.3. Mecanismos de adaptação, patogenicidade e virulência de <i>Xylella fastidiosa</i>	21
1.4. Análise de transcritomas através de microarranjos de DNA.....	23
1.5. A ação de antibióticos no perfil de expressão gênica global de bactérias	25
2. Objetivos	27
3. Procedimentos Experimentais	28
3.1. Manutenção e cultivo de <i>Xylella fastidiosa</i>	28
3.2. Ensaio de viabilidade celular	28
3.3. Tratamento de <i>Xylella fastidiosa</i> com dose subletal de estreptomicina	28
3.4. Isolamento de RNA total	29
3.5. Preparo de cDNA marcado e hibridização com os microarranjos.....	31
3.6. Análise dos dados de microarranjos	33
3.7. Transcrição reversa seguida de reação em cadeia quantitativa da polimerase em tempo real (RT-qPCR em tempo real)	34
3.8. Infecção de <i>Nicotiana clevelandii</i> com <i>X. fastidiosa</i>	36
3.9. Pirosequenciamento do genoma de <i>X. fastidiosa</i>	38
3.9.1. Purificação de DNA genômico de <i>Xylella fastidiosa</i>	38
3.9.2. Análise do DNA genômico das cepas 9a5c e J1a12.	38
3.9.3. Genotipagem de <i>Xylella fastidiosa</i>	41
3.9.4. Pirosequenciamento de DNA.....	43
3.9.10. Análises bioinformáticas	44
4. Resultados e Discussão	45
4.1. Análise do perfil de expressão gênica global de <i>Xylella fastidiosa</i> em resposta ao tratamento com dose subletal de estreptomicina	45
4.1.1 Análise do perfil de expressão gênica de CDS selecionadas em resposta ao tratamento com dose subletal de estreptomicina	56
4.1.2. Genes modulados no tratamento de <i>X. fastidiosa</i> com dose subletal de estreptomicina	58
4.2. Análise integrada de perfis de expressão gênica global de <i>Xylella fastidiosa</i>	65
4.3. Sequenciamento do genoma da cepa J1a12 de <i>Xylella fastidiosa</i>	69
4.3.1. Confirmação do fenótipo não-virulento da cepa J1a12.	69
4.3.2. Pirosequenciamento e montagem dos genomas das cepas 9a5c e J1a12.	72
4.3.3. Comparação preliminar do genoma de J1a12 com o genoma de 9a5c.....	75
5. Discussão Final	81
6. Referências Bibliográficas.....	84

Anexos.....	98
Tabelas Suplementares	98
Figura S1	107
Material Suplementar.....	118
Pirosequenciamento do DNA genômico de <i>X. fastidiosa</i>	118
Súmula Curricular	137

1. Introdução

1.1. O fitopatógeno *Xylella fastidiosa*

Xylella fastidiosa é uma bactéria Gram-negativa, não-flagelada, que coloniza o xilema de diversas espécies vegetais e é transmitida às plantas hospedeiras por insetos vetores da família Cicadellidae popularmente conhecidos como cigarrinhas. Estes insetos se alimentam do fluido que percorre o xilema da planta e são infectados ao se alimentarem da seiva de plantas contaminadas. As bactérias ficam restritas à parte anterior do intestino do inseto, aderidas ao forro cuticular do cibário, sendo transmitidas para plantas sadias em alimentações subsequentes. No hospedeiro vegetal, *X. fastidiosa* forma um biofilme compacto, que eventualmente ocasiona oclusão dos vasos do xilema (Purcell e Hopkins, 1996; Hopkins e Purcell, 2002; Alves *et al.*, 2003; Newman *et al.*, 2003; Redak *et al.*, 2004; Almeida *et al.*, 2005; Chatterjee *et al.*, 2008a). Imagens de agregados de células e do biofilme formados por *X. fastidiosa* na planta e no inseto-vetor estão apresentadas na Figura 1.

O gênero *Xylella* é composto por uma única espécie, *Xylella fastidiosa*, e faz parte da família *Xanthomonadaceae*, a qual inclui também os gêneros *Xanthomonas* e *Stenotrophomonas* (de Mello Varani *et al.*, 2009; Pieretti *et al.*, 2009). A família *Xanthomonadaceae* pertence à subdivisão gama das proteobactérias, que é modelo em diversificação bacteriana, incluindo espécies comensais e de vida-livre, simbioses intracelulares e patógenos de plantas e animais (Woese, 1987; Lerat *et al.*, 2005). Diversas espécies dos gêneros *Xylella* e *Xanthomonas* são fitopatogênicas ao contrário de algumas espécies de *Stenotrophomonas*, como *S. maltophilia*, que estabelecem interações benéficas com plantas (Chatterjee *et al.*, 2008a; Kay e Bonas, 2009; Ryan *et al.*, 2009). Por outro lado, *S. maltophilia* é um patógeno humano oportunista, com importância crescente em infecções hospitalares (Ryan *et al.*, 2009).

Análises filogenéticas multigênicas possibilitaram o agrupamento de isolados de *X. fastidiosa* em cinco subespécies: *piercei*, *sandyi*, *multiplex*, *pauca* e *tashke* as quais estão associadas a diferentes hospedeiros como esquematizado na Figura 2 (Rodrigues *et al.*, 2003; Schaad *et al.*, 2004; Schuenzel *et al.*, 2005; Randall *et al.*, 2009).

Em algumas espécies de plantas, *X. fastidiosa* existe como uma bactéria endofítica aparentemente benigna enquanto que nos hospedeiros suscetíveis, sua proliferação no xilema resulta em sintomas que incluem clorose, lesões necróticas nas folhas, nanismo, deterioração dos galhos, frutos não-adequados, e, eventualmente, morte da planta. Entre as doenças associadas à infecção por *X. fastidiosa*, estão a doença de Pierce (PD) das videiras, a escaldadura foliar de amendoeiras (*almond leaf scorch*, ALS), a escaldadura foliar do cafeeiro ou requeima do café (*coffee leaf scorch*, CLS), a escaldadura foliar de ameixeiras

(*plum leaf scald*, PLS) e a clorose variegada dos citros (CVC), popularmente conhecida como amarelinho (Purcell e Saunders, 1999; Purcell *et al.*, 1999; Hopkins e Purcell, 2002).

A doença de Pierce está presente no estado da Califórnia, nos Estados Unidos, desde 1884 e ocorre também em videiras de outros estados do sul dos EUA, como a Flórida e Texas, além do México e países da América Central. A partir de 1990, a doença agravou-se em razão do aumento de uma espécie de cigarrinhas oriunda do extremo sul dos EUA e México. Em 1999, mais de 300 acres de videiras da região de Temecula na Califórnia foram destruídos pela PD (Hopkins, 1989; Hopkins e Purcell, 2002). 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/>).

A CVC foi identificada pela primeira vez, em 1987, em laranjais do Estado de São Paulo (Rossetti *et al.*, 1990; Chang *et al.*, 1993), e desde então já foi detectada em pomares citrícolas de outras regiões do Brasil (<http://www.embrapa.gov.br>). Segundo levantamento realizado em julho de 2009, pelo Fundo de Defesa da Citricultura (<http://www.fundecitrus.com.br>), a CVC atinge 39% das plantas do parque citrícola do Estado de São Paulo, sendo que nos últimos 5 anos, observou-se uma queda da incidência da doença, possivelmente relacionada à produção de mudas sadias, aumento do uso de inseticidas para o controle do inseto vetor e a eliminação de pomares velhos, afetados pelo amarelinho.

Os sintomas da CVC incluem clorose intervenal, redução do tamanho das folhas, lesões com presença de goma na face abaxial das folhas, que eventualmente podem sofrer necrose. Com o avanço da CVC, pode ser observada a desfolha dos ramos mais altos da planta, enrijecimento das cascas dos frutos, os quais apresentam tamanho reduzido e assimetria (Figura 3). As características dos frutos de plantas doentes não somente reduz seu valor comercial *in natura*, como também prejudica a indústria de sucos (Neves *et al.*, 2001; de Souza *et al.*, 2005). Para comparação, apresentamos na Figura 4, o aspecto da folha e do fruto de videiras infectadas com *X. fastidiosa*, cepa Temecula.

Como já mencionado, as plantas infectadas por *X. fastidiosa* apresentam os vasos do xilema bloqueados por agregados celulares bacterianos, os quais são adensados pela presença do exopolissacarídeo (goma) secretado pela bactéria (Alves *et al.*, 2003; Newman *et al.*, 2004; Roper *et al.*, 2007a; Francis *et al.*, 2008). Assim, os sintomas das plantas infectadas por *X. fastidiosa* parecem estar relacionados aos estresses hídrico e/ou nutricional decorrentes da oclusão do xilema (Purcell e Hopkins, 1996; Hopkins e Purcell, 2002; Chatterjee *et al.*, 2008a). Contudo, muito possivelmente, os sintomas da doença também teriam associação com a resposta sistêmica do hospedeiro suscetível à interação com o patógeno (Jones e Dangl, 2006; Gambetta *et al.*, 2007; Lin *et al.*, 2007; Purcino *et al.*,

2007; Silva-Stenico *et al.*, 2009).

Múltiplos mecanismos de patogenicidade e determinantes de virulência foram sugeridos a partir da análise do repertório de genes codificados no genoma de *X. fastidiosa*, alguns destes já experimentalmente testados. Mais adiante este tópico será abordado em maior detalhe.

Figura 1. Imagens de agregados de células e do biofilme formados por *X. fastidiosa* na planta e no inseto-vetor. Em (A) e (B) são mostrados agregados bacterianos de *X. fastidiosa* cepa 9a5c em vasos do xilema de tabaco. Em (C), é mostrado o biofilme formado pela cepa Temecula-1 no aparelho digestivo do inseto vetor. As imagens foram extraídas de (Alves *et al.*, 2003; Newman *et al.*, 2004).

Figura 2. Árvore filogenética de isolados de *Xylella fastidiosa*. O filograma mostrado em (A) foi construído a partir de sequências do rRNA 16S e da região ITS (*Internal Transcribed Region*) do operon de rRNA de diversos isolados de *X. fastidiosa*. Os hospedeiros frequentemente associados as 5 subespécies estão indicados em (B). O filograma (A) foi extraído de (Randall *et al.*, 2009).

Figura 3. Sintomas típicos Clorose Variegada dos Citros (A) Frutos assimétricos e com tamanho reduzido e **(B)** folhas apresentando lesões cloróticas. Imagens extraídas de <http://www.fundecitrus.com.br>.

Figura 4. Sintomas típicos da Doença de Pierce em videiras. (A) Mudança de cor e ressecamento das folhas, concêntricos a partir das bordas e **(B)** frutos secos com aspecto de uva-passa. Imagens extraídas de <http://www-plb.ucdavis.edu/>

1.2. Genomas de *Xylella fastidiosa*

Após o trabalho pioneiro de sequenciamento e anotação do genoma da cepa 9a5c de *X. fastidiosa* (Simpson *et al.*, 2000) a qual foi isolada de laranjeiras e comprovadamente associada à CVC (Li *et al.*, 1999), genomas de outros 5 isolados/cepas foram completa ou parcialmente sequenciados (Tabela 1): cepa Temecula-1, isolada de videiras e agente etiológico da doença de Pierce; cepas Dixon, M12 e M23, associadas à escaldadura da amendoeira e cepa Ann-1 agente causal escaldadura da espírradeira (Simpson *et al.*, 2000; Bhattacharyya *et al.*, 2002a; Bhattacharyya *et al.*, 2002b; Van Sluys *et al.*, 2003; Chen *et al.*, 2007). Além disso, recentemente foi realizada a re-anotação de quatro destes genomas (9a5c, Temecula-1, Dixon e Ann-1) (de Mello Varani *et al.*, 2008; de Mello Varani *et al.*, 2009), a qual está disponível em <http://www.xylella.lncc.br>.

Tabela 1. Genomas sequenciados de cepas de *Xylella fastidiosa*

Cepa	Doença	Replicons	pb	Número de genes/CDS ^e	% A+T	rRNA 5S-16S-23S
9a5c ^a	CVC	3	2 679 305	2 782 (2 430) ^f	47,33	2
			51 158	65 (61) ^f	50,43	
			1 285	2 (2) ^f	44,4	
Ann-1 ^b	OLS	2	2 625 581	2 870	48,1	2
			30 270	36	51	
Dixon ^b	ALS	1	2 434 849	2 681	48	2
M12 ^c	ALS	1	2 475 130	2 104	48,1	2
M23 ^c	ALS	2	2 535 690	2 261	48,24	2
			38 297	40	50,08	
Temecula ^d	PD	2	2 519 802	2 034	48,22	2
			1 345	2	46,21	

^a Genoma completo (Simpson *et al.*, 2000)

^b Genomas incompletos (Bhattacharyya *et al.*, 2002b)

^c Genomas incompletos (Chen *et al.*, 2007)

^d Genoma completo (Van Sluys *et al.*, 2003)

^e A diferença entre o número de genes reflete critérios de anotação distintos em cada caso.

^f O número de genes originalmente anotados por (Simpson *et al.*, 2000) foi revisto após re-anotação completa do genoma da 9a5c (de Mello Varani *et al.*, 2009) e está indicado entre parênteses.

Análises comparativas das sequências e dos dados de anotação dos genomas destas cepas revelaram expressiva variabilidade, o que corrobora com a separação em distintos grupos ou subespécies apresentada na Figura 2. Foi verificado que uma região da cepa Dixon está ausente nos genomas das cepas 9a5c e Ann-1, a qual potencialmente está

associada à especificidade ao hospedeiro. Observação semelhante foi feita sobre a cepa Temecula que também apresenta região exclusiva (Bhattacharyya *et al.*, 2002a; Bhattacharyya *et al.*, 2002b; Van Sluys *et al.*, 2002; Nunes *et al.*, 2003; Van Sluys *et al.*, 2003; Moreira *et al.*, 2005; de Mello Varani *et al.*, 2008; de Mello Varani *et al.*, 2009).

Xylella fastidiosa pertence a um grupo de organismos com genomas de tamanho relativamente pequeno e pode-se observar que os prováveis eventos de deleção genômica parecem ter sido mais abundantes do que os eventos de ganho por duplicação ou transferência lateral, ainda que sequências de fagos tenham se inserido em seu genoma (Van Sluys *et al.*, 2003; de Mello Varani *et al.*, 2008; de Mello Varani *et al.*, 2009; Pieretti *et al.*, 2009). A ausência de genes e vias importantes nesta espécie, decorrente desta evolução convergente de redução de seu genoma, provavelmente foi favorecida pela adaptação destas bactérias ao crescimento no restrito ambiente do xilema e a dependência de um vetor animal para sua propagação (Van Sluys *et al.*, 2002; Chatterjee *et al.*, 2008a; Pieretti *et al.*, 2009).

A comparação de genomas completos das cepas de *X. fastidiosa*, também revelou a ocorrência de rearranjos genômicos, aparentemente decorrentes da presença de genes de prófagos, especificamente integrases, localizados justamente nas bordas destes rearranjos ou mesmo flanqueando prováveis regiões de inserção e deleção. Estes dados evidenciam o impacto de regiões de prófagos na organização e diferenciação destes genomas (Van Sluys *et al.*, 2002; de Mello Varani *et al.*, 2008; de Mello Varani *et al.*, 2009).

1.3. Mecanismos de adaptação, patogenicidade e virulência de *Xylella fastidiosa*

A análise do repertório de genes previamente anotados no genoma da cepa 9a5c (Simpson *et al.*, 2000) seguida por estudos comparativos *in silico* dos genomas de diferentes cepas de *X. fastidiosa* possibilitou a identificação de vários genes potencialmente relacionados a processos de adaptação, patogenicidade e virulência nesta bactéria (Bhattacharyya *et al.*, 2002a; Bhattacharyya *et al.*, 2002b; Van Sluys *et al.*, 2003; Setubal *et al.*, 2005; Doddapaneni *et al.*, 2006; Chatterjee *et al.*, 2008a; de Mello Varani *et al.*, 2009). Outras abordagens, tais como análise de polimorfismos genéticos ou comparação de genomas através de microarranjos de DNA têm contribuído para revelar variações na composição gênica entre cepas ou isolados de *Xylella fastidiosa* (Chen *et al.*, 1992; Pooler *et al.*, 1996; Rosato *et al.*, 1998; Chen *et al.*, 2000; da Costa *et al.*, 2000; Mehta *et al.*, 2001; Costa de Oliveira *et al.*, 2002; Coletta-Filho e Machado, 2003; Nunes *et al.*, 2003; Koide *et al.*, 2004b; Scally *et al.*, 2005; Schuenzel *et al.*, 2005; da Silva *et al.*, 2007; Wickert *et al.*, 2007; Almeida *et al.*, 2008).

A estes estudos somam-se análises do transcrito de *X. fastidiosa* que

evidenciaram vários genes que estão potencialmente envolvidos na patogenicidade, virulência, competição e sobrevivência deste fitopatógeno (de Souza *et al.*, 2003; de Souza *et al.*, 2004; de Souza *et al.*, 2005; Pashalidis *et al.*, 2005; Koide, 2006; Koide *et al.*, 2006b; da Silva Neto *et al.*, 2007; Shi *et al.*, 2007; da Silva Neto *et al.*, 2008; Zaini *et al.*, 2008; Killiny e Almeida, 2009a; Shi *et al.*, 2009a; Shi *et al.*, 2009b; Traversolo *et al.*, 2009; Fogaça *et al.*, 2010).

Estudos em escala genômica conjuntamente aos estudos que utilizam abordagens de genética molecular, como análise de mutantes, e também de biologia celular têm corroborado a hipótese de que patogenicidade e virulência de *X. fastidiosa* compreendem múltiplos fatores, entre estes, como já mencionamos, está à formação do biofilme que promoveria a oclusão do xilema e consequente estresse hídrico (McElrone *et al.*, 2001; McElrone *et al.*, 2003; Newman *et al.*, 2003; Scarpari *et al.*, 2003; Osiro *et al.*, 2004; de Souza *et al.*, 2005; Guilhabert e Kirkpatrick, 2005; Meng *et al.*, 2005; Li *et al.*, 2007; Shi *et al.*, 2007; Chatterjee *et al.*, 2008a; Chatterjee *et al.*, 2008b; McElrone *et al.*, 2008).

Outros fatores possivelmente relevantes para o estabelecimento da infecção por *Xylella fastidiosa* incluem:

(i) migração da bactéria ao longo do vaso do xilema pelo movimento do tipo *twitching* proporcionado pela presença do fluxo de líquido e pela presença de *pili* longos (pilus tipo-IV) (Meng *et al.*, 2005; Chatterjee *et al.*, 2008a; Chatterjee *et al.*, 2008c; De La Fuente *et al.*, 2008);

(ii) migração da bactéria entre os vasos do xilema proporcionada pela expressão de enzimas capazes de degradar a parede que separa os pontos de comunicação entre os vasos do xilema (Newman *et al.*, 2003; Scarpari *et al.*, 2003; Koide *et al.*, 2004b; Newman *et al.*, 2004; Koide *et al.*, 2006b; Maria Fedatto *et al.*, 2006; Nogaroto *et al.*, 2006; Roper *et al.*, 2007b; Chatterjee *et al.*, 2008a; Warren e Kirkpatrick, 2009);

(iii) captação de ferro e outros metais de transição pela *X. fastidiosa* causando sua depleção no xilema das plantas infectadas (Simpson *et al.*, 2000; Meidanis *et al.*, 2002; Silva-Stenico *et al.*, 2005; Zaini *et al.*, 2008);

(iv) secreção de drogas e toxinas que facilitem à sobrevivência da bactéria no xilema (Simpson *et al.*, 2000; Meidanis *et al.*, 2002; Koide *et al.*, 2004b; de Souza *et al.*, 2005; Pashalidis *et al.*, 2005; Zaini *et al.*, 2008);

(v) mudanças fenotípicas na bactéria que facilitem sua transmissão pelo vetor e/ou sua disseminação sistêmica no hospedeiro vegetal (Guilhabert e Kirkpatrick, 2005; Killiny e Almeida, 2008; Killiny e Almeida, 2009b; Killiny e Almeida, 2009a).

É importante ressaltar que vários destes fatores são controlados pelo acúmulo de DSF (*Diffusible Signaling Factor*) uma pequena molécula sinalizadora (ácido 12-metil-tetradecanóico ou ácido cis-11-metil-2-dodecenóico) cuja síntese acompanha o aumento da

densidade celular, como ocorre com diversos outros fitopatógenos (da Silva *et al.*, 2001; Scarpari *et al.*, 2003; Wang *et al.*, 2004; Colnaghi Simionato *et al.*, 2007; Chatterjee *et al.*, 2008c; Dow, 2008; Newman *et al.*, 2008).

Os aspectos que abordamos demonstram que, dez anos após a divulgação do sequenciamento completo do genoma da cepa 9a5c de *X. fastidiosa* (Simpson *et al.*, 2000), muito se avançou na compreensão dos processos envolvidos na colonização dos vasos do xilema de plantas suscetíveis e também do inseto vetor. Ainda assim, podemos presumir a existência de mecanismos e fatores adicionais potencialmente importantes para colonização específica e seletiva de hospedeiros vegetais e insetos por este fitopatógeno, os quais permanecem desconhecidos ou pouco explorados.

1.4. Análise de transcritomas através de microarranjos de DNA

Nos últimos anos, temos acompanhado um aumento massivo na quantidade de informação acumulada em sequências de DNA. No entanto, os bilhões de bases das sequências de DNA não nos indicam a função de todos os genes ou como as células trabalham e formam organismos. O objetivo de estudos genômicos não é só simplesmente produzir um catálogo com o nome dos genes e suas funções, mas principalmente, entender como os componentes desse quebra-cabeça trabalham juntos e assim compreender o funcionamento de células e organismos. Para que possamos aproveitar e explorar ao máximo toda essa informação gerada pelo sequenciamento de DNA, novas tecnologias são então necessárias.

Entre as mais poderosas e versáteis ferramentas para estudos de genômica funcional estão os microarranjos (*microarrays*) de DNA. Os microarranjos são preparados pela deposição, em lâminas de vidro, de moléculas de DNA geradas por PCR específicas para cada segmento gênico diretamente do DNA genômico de interesse ou por PCR de clones de bibliotecas de cDNA. Alternativamente, são utilizados oligonucleotídeos (*oligoarrays*) com tamanho variando de 20 a 70 nucleotídeos (*oligoarrays*). Em ambos os casos, os fragmentos de DNA imobilizados na lâmina de vidro, em arranjo precisamente ordenado em alta densidade, representam milhares ou dezenas de milhares de genes (Lockhart e Winzeler, 2000; Marques e da Silva, 2004; Venkatasubbarao, 2004; Hoheisel, 2006). Hoje em dia, além da possibilidade da confecção “doméstica” de microarranjos ou de *oligoarrays*, há diversas plataformas comercialmente disponíveis (ex: Agilent, Affymetrix, NimbleGen). Além disso, a tecnologia evoluiu para obtenção dos chamados *tiling* (*oligo*)*arrays* que representam cromossomos completos, permitindo também a avaliação da expressão de regiões genômicas não codificadoras de proteínas (Bertone *et al.*, 2004; Mockler *et al.*, 2005).

Os microarranjos de DNA permitem, com relativa facilidade, avaliar simultaneamente

e quantitativamente os níveis de expressão de milhares de genes (abundância de transcritos) de células submetidas a situações ou tratamentos diversos (Brazma e Vilo, 2000; Hoheisel, 2006). O RNA total ou mRNA é primeiramente isolado de diferentes tecidos ou de células submetidas a tratamentos apropriados, para ser em seguida utilizado como molde para a síntese de cDNA, que será marcado por fluoróforos antes de ser hibridizado aos microarranjos. A intensidade do sinal fluorescente resultante da hibridização do cDNA a cada ponto (DNA depositado) do microarranjo é quantificada e algoritmos de normalização para correção de fatores, tais como diferenças de intensidades de incorporação e variações topográficas na lâmina, são aplicados aos valores obtidos. O cálculo da razão das intensidades de dois sinais fluorescentes para um mesmo ponto no microarranjo reflete a expressão relativa do gene ali representado, entre as duas amostras biológicas que estão sendo comparadas e cujos cDNAs foram marcados com fluoróforos que emitem sinal fluorescente em comprimentos de onda distintos. Deste modo, a razão das intensidades dos dois sinais fluorescentes em cada ponto permite avaliar a expressão relativa de cada gene representado no microarranjo, ou seja, se o gene em questão está mais ou se está menos expresso nas amostras biológicas que estão sendo comparadas. Genes com perfis de expressão similares podem ser agrupados (*clustering*) e relações de parentesco entre as amostras ou entre os genes podem ser estabelecidas através da aplicação de metodologias estatísticas, tais como o agrupamento hierárquico (Eisen *et al.*, 1998; Holloway *et al.*, 2002; Quackenbush, 2002). Após a prospecção inicial com o uso da tecnologia de microarranjos de DNA, que aponta conjuntos selecionados e relevantes de genes que apresentam expressão diferencial, procede-se a validação através de técnicas como RT-qPCR (PCR quantitativo precedido de transcrição reversa).

Através de experimentos de hibridização de microarranjos de DNA, é possível o monitoramento da expressão gênica para conhecermos onde e/ou quando um gene é expresso, uma informação que pode embasar o entendimento do papel desempenhado pela proteína codificada por tal gene (Lockhart e Winzeler, 2000). Desta forma, a análise de perfis de expressão de genes (transcritomas) de células expostas a diversas situações é uma etapa importante para a compreensão dos mecanismos envolvidos nos mais diferentes processos celulares incluindo, por exemplo, a resposta das células à ação de drogas ou a estresses ambientais variados.

Entre as várias aplicações dos microarranjos de DNA, esta metodologia tem sido muito utilizada para a detecção de alterações na expressão gênica de bactérias frente a diferentes estímulos, uma vez que a regulação ao nível da transcrição é preponderante em procariotos (Lucchini *et al.*, 2001; Schoolnik, 2002a; Schoolnik, 2002b; Conway e Schoolnik, 2003). Esta tecnologia tem possibilitado a elucidação de circuitos regulatórios em diversas bactérias e também a definição de diversos *regulons* (Laub *et al.*, 2000; DeLisa *et al.*, 2001;

Baichoo *et al.*, 2002; Koide *et al.*, 2006b; Michel *et al.*, 2006; Deslandes *et al.*, 2007; Maurer *et al.*, 2007; Zaini *et al.*, 2008; Schmid *et al.*, 2009).

A análise do perfil global de expressão gênica através de microarranjos de DNA tem sido também importante na elucidação de mecanismos associados à patogenicidade e virulência em bactérias. Por exemplo, microarranjos de DNA de *Mycobacterium tuberculosis* revelaram o circuito genético controlado por *sigH*, um regulador transcricional que quando mutado altera a virulência desta bactéria (Kaushal *et al.*, 2002). Além dos trabalhos que investigaram o transcriptoma de *X. fastidiosa* e que apontaram para genes associados a patogenicidade, os quais foram anteriormente citados, a literatura é repleta de exemplos da aplicação desta tecnologia na elucidação de redes genéticas associadas a virulência de patógenos bacterianos (Schoolnik, 2002b; Burrack e Higgins, 2007).

1.5. A ação de antibióticos no perfil de expressão gênica global de bactérias

A maioria dos antibióticos utilizados na terapia de doenças infecciosas deriva de moléculas produzidas por micróbios, como fungos e bactérias, as quais, provavelmente, desempenham algum papel nos fenômenos de interação microbiana que ocorrem em seu habitat natural. Assume-se, assim, que os antibióticos tenham atividades biológicas adicionais, além da atividade inibitória, podendo eventualmente atuar como moléculas sinalizadoras em uma comunidade microbiana (Goh *et al.*, 2002; Hoffman *et al.*, 2005; Linares *et al.*, 2006).

A avaliação do perfil de expressão gênica global de microorganismos expostos a diferentes concentrações de tipos distintos de antibióticos tem sido empregada na investigação de seus respectivos mecanismos de ação e também na identificação das respostas fisiológicas resultantes da exposição ao antibiótico, incluindo mecanismos de resistência e de patogenicidade (Wolf *et al.*, 2000; Evers *et al.*, 2001; Fitzgerald *et al.*, 2001; Betts *et al.*, 2003; Shaw *et al.*, 2003; Hansen *et al.*, 2004; Brazas e Hancock, 2005; Qiu *et al.*, 2005; Overhage *et al.*, 2008; Jones *et al.*, 2009).

Uma abordagem, bastante frequente é o estudo dos efeitos de concentrações subletais de agentes antimicrobianos, tanto os antibióticos não-peptídicos como os peptídeos antimicrobianos (AMP), na expressão gênica. O interesse nesta abordagem decorre do fato de que no tratamento com doses letais do antibiótico, eventualmente uma fração da população bacteriana poderá ficar exposta a doses subletais. As evidências de que as bactérias respondem a doses subletais alterando seu perfil de expressão gênica são claras. Por exemplo, o tratamento de *E. coli* com doses subletais de cecropina demonstrou que esse AMP altera os níveis transcricionais de 26 genes (Hong *et al.*, 2003). Outro estudo demonstrou que o crescimento de *Salmonella typhimurium* na presença de doses subletais de AMPs ativa regulons de virulência, como o sistema PhoP/PhoQ (Groisman *et al.*, 1989;

Ernst *et al.*, 2001), que são responsáveis pela repressão da transcrição de genes necessários para síntese de flagelo e para invasão do hospedeiro associada ao sistema de secreção tipo III (Bader *et al.*, 2003).

Quando tratada com dose subletal de estreptomicina, tetraciclina, puromicina, cloranfenicol ou eritromicina, a bactéria *Streptococcus pneumoniae* R6, apresentou perfis de expressão gênica consistentes com o modo de ação desses inibidores traducionais. Em resposta ao tratamento, foi observada a alteração nos níveis de transcritos de genes que codificam proteínas ribossomais e fatores de tradução, dentre vários outros genes relacionados ao processo de tradução, além de genes relacionados à resposta de choque térmico (Ng *et al.*, 2003). Interessantemente, a exposição de *Pseudomonas aeruginosa* a doses subletais de tobramicina, tetraciclina ou norfloxacinina induz a formação de biofilme, em decorrência da alteração dos níveis de di-GMP cíclico, e promove também a expressão de determinantes de virulência (Hoffman *et al.*, 2005; Linares *et al.*, 2006). Por outro lado, quando exposta ao peptídeo antimicrobiano LL-37, observa-se supressão da expressão de genes essenciais para formação do biofilme (Overhage *et al.*, 2008). A bactéria *Salmonella typhimurium* em resposta à dose subletal de eritromicina (inibidor da tradução) ou rifampicina (inibidor da transcrição) modula a expressão de 5% de seus genes (Goh *et al.*, 2002).

Além das aplicações na medicina humana e veterinária, os antibióticos são também utilizados na agricultura, principalmente estreptomicina e oxitetraciclina (Vidaver, 2002). A estreptomicina é um aminoglicosídeo solúvel em água, comercializado como um sal de sulfato ou nitrato de estreptomicina, em formulação conjunta com oxitetraciclina (Agri-micina) ou não (Agri-strep). Desde sua introdução em 1955, este antibiótico tem sido o mais usado na agricultura contra diversos patógenos tais como *Erwinia amylovora*, *Pseudomonas cichorii* e várias cepas de *Pseudomonas syringae*, *Xanthomonas campestris* pv. *vesicatoria*, *Agrobacterium tumefaciens*, dentre outros. Entretanto, 5 anos após sua introdução na agricultura, já surgiram fitopatógenos mutantes resistentes a estreptomicina (McManus *et al.*, 2002; Vidaver, 2002).

Como já mencionado, a estreptomicina é um inibidor da síntese de proteínas, que resulta na perda de acurácia do processo de tradução e em um acúmulo de polipeptídeos com erros. Já foi demonstrado que a estreptomicina interage diretamente com o centro de descodificação da subunidade menor do ribossomo (sítio A), através de interação com resíduos conservados do rRNA 16S e com a proteína ribossomal S12 (Carter *et al.*, 2000; Ogle *et al.*, 2003). A maioria das cepas resistentes à estreptomicina apresentam mutações no gene *rpsL* que codifica a proteína S12, as quais estabilizam a conformação da subunidade 30S, restringindo a taxa de erros de tradução. (Ogle *et al.*, 2003; Holberger e Hayes, 2009).

2. Objetivos

O objetivo inicial deste trabalho foi analisar, utilizando microarranjos de DNA, o perfil de expressão gênica global de *Xylella fastidiosa* em resposta ao tratamento com dose subletal de estreptomicina, como uma estratégia para identificar suas respostas fisiológicas decorrentes da exposição à dose subinibitória deste antibiótico. Como fundamentado na Introdução, em concentrações subinibitórias, agentes antimicrobianos modulam a expressão gênica bacteriana, sendo que o conjunto de genes que é modulado depende tanto da cepa bacteriana e como da natureza do agente antimicrobiano, podendo apontar para determinantes de virulência e de resistência.

Propusemos também comparar o perfil de expressão gênica de *X. fastidiosa* em resposta ao tratamento com estreptomicina com perfis de expressão observados em outras condições experimentais, cujos dados estão publicamente disponíveis. A análise integrada de dados possibilita a prospecção de correlações entre genes modulados em mais de uma condição testada, que poderão ser exploradas futuramente visando uma caracterização mais detalhada de regulação de expressão de genes selecionados e/ou a identificação de alvos para análise funcional.

Finalmente, incorporamos um objetivo adicional ao nosso trabalho, onde propusemos a realização do sequenciamento do genoma da cepa J1a12 de *X. fastidiosa*, utilizando a tecnologia de pirosequenciamento de DNA no equipamento 454 Genome Sequencer-FLX, instalado no Centro Avançado de Tecnologias em Genômica (CATG) do Departamento de Bioquímica do IQUSP. A cepa J1a12 foi originalmente isolada de laranjeiras com sintomas de CVC, e até o presente, dentre as cepas isoladas de citros, é a única cepa comprovadamente suscetível a manipulação genética (Monteiro *et al.*, 2001b; da Silva Neto *et al.*, 2002; Koide *et al.*, 2004a). Entretanto, foi demonstrado que esta cepa exibe fenótipo menos virulento em *Nicotiana tabacum* e *Citrus sinensis* em comparação à cepa 9a5c (Koide *et al.*, 2004b). A comparação entre genomas das cepas J1a12 e 9a5c utilizando microarranjos de DNA revelou diferenças significativas que podem estar correlacionadas com o fenótipo menos virulento da cepa J1a12 (Koide *et al.*, 2004b). Assim, o sequenciamento do genoma da J1a12 que propusemos, visa aprofundar a comparação dos genomas das cepas J1a12 e 9a5c com o propósito de confirmarmos as diferenças preditas pela análise de microarranjos bem como identificarmos diferenças adicionais que contribuam para explicar as características fenotípicas da cepa J1a12.

3. Procedimentos Experimentais

3.1. Manutenção e cultivo de *Xylella fastidiosa*

As cepas 9a5c (Li *et al.*, 1999) e J1a12 (Monteiro *et al.*, 2001b) de *X. fastidiosa* foram cedidas pelo Fundecitrus (Araraquara, SP) e seu cultivo foi realizado em meio complexo PW (Davis *et al.*, 1981) modificado (PWG). O meio PWG é constituído de fitona peptona 4g/L; peptona de caseína com digestão tríptica 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%. A manutenção das culturas foi feita placas de meio PWG-ágar (PWG-contendo ágar 1,5%) com repiques semanais e incubação a 28°C, sendo o número de passagens controlado. Após 30 passagens, novo cultivo era iniciado a partir de estoques congelados a -80°C. Para o cultivo em meio líquido, as bactérias após 7 dias de cultivo em meio sólido eram transferidas para 50 mL de meio PWG e mantidas a 28°C com 100 rpm de agitação por mais 7 dias. Após esse período, eram adicionados 50 mL de meio PWG fresco às culturas que permaneciam a 28°C e 100 rpm, por mais 2 dias. Estoques de células (diversas passagens) foram mantidos em PWG contendo 50% de glicerol em freezer a -80°C.

3.2. Ensaio de viabilidade celular

Células de *X. fastidiosa* 9a5c previamente crescidas em meio 100 mL de meio PWG por 14 dias foram coletadas e ressuspensas em 20mL de meio PWG fresco ($DO_{600nm} \cong 0,400$) e alíquotas de 90µL da suspensão bacteriana foram incubadas, em triplicatas, com diferentes concentrações de sulfato de estreptomicina (0,1 µg/mL a 100 µg/mL) ou água como controle a 28°C por 15min, 30min, 60 min, 3h e 18h. Ao final de cada tempo de incubação, as células foram misturadas a 900µL de meio PWG fresco. Após medição da DO_{600nm} , alíquotas de 20µL de células expostas ou não às diferentes concentrações de antibióticos pelos tempos indicados foram coletadas e transferidas para placas de meio PWG-ágar, seguindo-se incubação a 28°C por 7 dias.

3.3. Tratamento de *Xylella fastidiosa* com dose subletal de estreptomicina

Após ~ 9 dias de cultivo, 100-200 mL de cultura da cepa 9a5c em meio PWG, foram divididos em alíquotas de 50 mL cada, sendo que essas alíquotas foram concentradas para um volume final de 10 mL antes da adição do antibiótico (10 µg/mL de estreptomicina) ou volume equivalente de água como controle, seguindo incubação a 28°C por 15, 30 e 60 minutos. Após o término de cada tempo de tratamento, alíquotas de 1 mL foram coletadas em duplicata ou triplicata. As amostras foram então centrifugadas a 3.200xg, por 5 min, o

sobrenadante foi descartado e o precipitado de células, armazenado a -80°C até que se procedesse o isolamento do RNA total. Foram também coletadas amostras antes da adição do antibiótico (tempo 0h) para ensaios de hibridização homotípica (ver adiante).

3.4. Isolamento de RNA total

RNA total foi extraído a partir das suspensões bacterianas tratadas ou não com estreptomicina, utilizando-se o reagente TRIZOL (Invitrogen) segundo o procedimento descrito pelo fabricante, tratado com DNase RQ1 (Promega) e em seguida, purificado utilizando-se fenol/clorofórmio. Para tal, os precipitados de células obtidos a partir de 1 mL das culturas foram suspensos com 1 mL de TRIZOL e incubados à temperatura ambiente por 5 minutos, seguindo-se adição de 200 μL de clorofórmio e incubação por mais 3 min à temperatura ambiente. As suspensões foram então centrifugadas a 16.000 $\times g$ por 15 minutos a 4°C . A fase aquosa foi recolhida em tubos limpos, adicionando-se em seguida 500 μL de isopropanol. Após incubação à temperatura ambiente por 15 min, as amostras foram submetidas a centrifugação a 16.000 $\times g$ por 10 min a 4°C . A fase aquosa foi recolhida em tubos limpos e a ela foi adicionado 1 mL de etanol 70%. As amostras foram centrifugadas a 11.000 $\times g$ por 5 min a 4°C , o etanol foi descartado e os precipitados foram secos em fluxo laminar antes de serem ressuspensos em 60 μL de água livre de Rnase (água pré-tratada com DEPC (dietilpirocarbonato) 0,1% como descrito em (Ausubel e Frederick, 1995)). Para a solubilização do RNA, as amostras foram aquecidas a 65°C por 10 min.

A quantificação da concentração de RNA foi feita através da determinação das absorvâncias (Abs) a 260nm e 280nm em espectrofotômetro ND-1000 (NanoDrop Technologies), sendo que $\text{Abs}_{260\text{nm}} = 1$ corresponde a 40 μg de RNA em 1 mL de solução quando a razão $\text{Abs}_{260\text{nm}}/\text{Abs}_{280\text{nm}} \cong 1,9$. A integridade do RNA total foi avaliada pela observação da proporção das bandas referentes aos rRNA 16S e 23S em eletroforese em gel de agarose 1,2% contendo formaldeído (Ausubel e Frederick, 1995), como exemplificado na Figura 4.

Em algumas amostras, a integridade do RNA foi avaliada no equipamento BioAnalyzer (Agilent) como mostrado na Figura 5, que permite a avaliação de amostras com concentrações de RNA entre 0,5–50 ng/ μL . A matriz do gel aplicada nos capilares do *chip* utilizado no BioAnalyzer contém uma mistura de fluoróforos e marcadores de peso molecular. A ligação dos fluoróforos ao marcador de tamanho e ao RNA resulta na emissão de fluorescência que é quantificada, permitindo a separação do rRNA 16S e 23S e a verificação da integridade do RNA total a partir da atribuição de um número de integridade do RNA (RIN) que varia de 1 a 10.

Finalmente, o RNA total foi tratado com RQ1 DNase (Promega) para a remoção de

eventual DNA genômico contaminante. Para cada 5µg de RNA eram adicionadas 5 U de RQ1 Dnase e a reação era incubada a 37°C por 30 minutos. Após a inativação da RQ1 Dnase, o RNA foi extraído com 500µL fenol/clorofórmio pH 5,2. A fase aquosa obtida após centrifugação a 5.000 xg por 5 minutos a temperatura ambiente foi transferida para um tubo limpo, ao qual foi acrescentado 500 µL de clorofórmio. Após centrifugação a 5.000 xg por 5 min, a fase aquosa foi removida e o RNA foi precipitado pela adição de 50 µL de acetato de sódio 3M, pH 5,2 e 1mL de etanol. O sedimento (RNA) obtido após centrifugação a 10.000xg por 10 min foi lavado com etanol 70%, centrifugando-se a 10.000xg por 10 minutos. O etanol foi descartado e o sedimento foi seco em fluxo laminar. Após solubilização do RNA em 50µL água livre de Rnase, a quantificação foi feita através da determinação das absorbâncias a 260nm e 280nm em espectrofotômetro ND-1000 (NanoDrop Technologies) e a integridade do RNA total purificado foi novamente avaliada no BioAnalyzer (Agilent).

Figura 5. Avaliação da integridade das amostras de RNA por eletroforese em gel de agarose em condições desnaturantes. As amostras de RNA total purificado de células não tratadas (canaletas 1 a 4) ou tratadas com estreptomicina por 15, 30 e 60 minutos (canaletas 5, 6 e 7, respectivamente) foi submetido a eletroforese em gel de agarose 1,2% contendo formaldeído, seguindo-se detecção com brometo de etídio. As bandas correspondem ao rRNA 23 S e 16 S (de cima para baixo).

Figura 6. Avaliação da integridade do RNA total no Bioanalyzer (Agilent Technologies).

Em **A**, está representada a separação das amostras de RNA total por eletroforese no BioAnalyzer. O canal L corresponde ao marcador de tamanho e os canais 1 e 2, correspondem, respectivamente, a amostras de RNA total de células não tratadas e tratadas com estreptomicina, ambas incubadas pelo tempo de 60 minutos. Podem ser observadas duas bandas mais intensas nos canais 1 e 2, as quais correspondem ao rRNA 16S e 23S. Em **B**, estão representados dois gráficos, que correspondem à emissão de fluorescência medida durante a eletroforese das amostras, onde o eixo y exibe as unidades de fluorescência medidas e o eixo x, o tempo de corrida em segundos. Podemos observar dois picos maiores que correspondem à fluorescência medida para o rRNA 16S e 23S, da esquerda para direita. O pico menor corresponde à fluorescência medida para o marcador de tamanho. O número de integridade do RNA (RIN) é bastante satisfatório.

3.5. Preparo de cDNA marcado e hibridização com os microarranjos

O cDNA fita simples marcado fluorescentemente foi gerado com o kit SuperScript™ Indirect cDNA Labeling System (Invitrogen) de acordo com as instruções do fabricante, a partir de 40 µg de RNA total na presença de 5 µg de iniciadores nonaméricos aleatórios. Após a síntese, híbridos de RNA/cDNA foram desnaturados por 10 minutos a 70°C e o RNA hidrolisado com NaOH. Em seguida foi adicionado 120 µL de *filter binding buffer* (5,3 M guanidina HCl e 150 mM CH₃COOK pH 4,8) a cada reação e essas foram transferidas para uma placa de purificação de 96 poços (Millipore Multiscreen *filter plate*). Após centrifugação

a 1250 xg por 1 minuto a 25°C, o filtrado foi descartado e o poço lavado com 200 µL de etanol 80% por 4 vezes, centrifugando entre cada lavagem por 1 minuto a 25°C. Uma última centrifugação adicional por 3 minutos foi feita para secar completamente os filtros. As amostras de cDNA foram eluídas duas vezes com 50 µL de 10 mM Tris pH 8 pré-aquecido a ~50°C, centrifugando por 2 minutos a 1.250 xg. O cDNA sintetizado em cada amostra foi quantificado, medindo-se Absorbância em 260nm e 280 nm. Em seguida as amostras foram secas sob centrifugação a vácuo e suspensas para a reação de ligação do cDNA ao fluoróforo, de acordo com as instruções do fabricante. Após uma nova etapa de purificação, o cDNA recuperado em cada amostra foi quantificado ($Abs_{260/280}$) e o fluoróforo total incorporado em cada amostra foi quantificado medindo-se Abs em 550 nm (Cy3 ou AF555) e em 650 nm (Cy5 ou AF647). As amostras foram secas novamente e suspensas em água previamente tratada com DEPC 0,1%.

Hibridizações na presença de 50% formamida e tampão de hibridização (GE Healthcare) foram realizadas por aproximadamente 16 horas em câmaras umedecidas e aquecidas a 42°C, utilizando-se microarranjos em lâmina de vidro que representam 2608 sequências codificadoras (CDS) de *Xylella fastidiosa* cepa 9a5c e que preparamos no laboratório de microarranjos do Departamento de Bioquímica do IQUSP como previamente descrito (Koide *et al.*, 2004b; Zaini, 2007; Zaini *et al.*, 2008). Breve descrição do microarranjo está apresentada abaixo. Para remover hibridizações inespecíficas, as lâminas foram lavadas com SSC 1x/SDS 0,2% por 10 minutos a 55°C; duas vezes com SSC 0,1x/SDS 0,2% também por 10 minutos a 55°C cada; SSC 0,1x por 10 minutos à temperatura ambiente e finalmente mergulhada 3-4 vezes em água deionizada. Todas as soluções de lavagem foram filtradas em membranas com poros de 0,22 µm. Em seguida as lâminas foram secas com gás nitrogênio e analisadas no *Generation III Scanner*TM (GE Healthcare) para obtenção das intensidades de fluorescências dos pontos dos microarranjos.

Os microarranjos foram preparados pela deposição robotizada de amostras de DNA em lâminas espelhadas tipo 7 (GE Healthcare), usando o equipamento *Generation III Microarray Spotter* (GE Healthcare). As amostras consistiam de fragmentos de DNA fita dupla de 250-1000 pb CDS-específicos previamente obtidos por amplificação por PCR do DNA genômico ou dos plasmídeos pXF51 e pXF1.5 isolados de *X. fastidiosa* cepa 9a5c com pares de oligonucleotídeos específicos para CDS anotadas no genoma desta cepa (Koide *et al.*, 2004b; Zaini, 2007; Zaini *et al.*, 2008). Este *spotter* permite a deposição de até 4608 amostras de DNA organizadas em 12 subconjuntos (*subarrays*) de 384 pontos (12 linhas x 32 colunas). O conjunto total de 4608 amostras é depositado em duplicada nas duas metades longitudinais da lâmina, configurando o que chamamos de *set A* e *set B*, os quais são considerados como réplicas técnicas. Após a deposição dos fragmentos de DNA, as lâminas foram submetidas a 50 mJ de luz UV, para que fitas de DNA sejam covalentemente

fixadas. As lâminas assim obtidas foram armazenadas em dessecador (umidade relativa ~15%) por no máximo seis meses.

3.6. Análise dos dados de microarranjos

Após a varredura dos microarranjos, as intensidades brutas de sinal das imagens foram extraídas usando o programa *Array Vision* 6.0 (Image Research/ Molecular Dynamics). Após a subtração do ruído local de cada ponto, a densidade de fluorescência média removida de artefatos (MTM Dens) foi coletada para cada ponto. Esta ferramenta do *Array Vision* permite a exclusão de *pixels* dentro do ponto que mostrem intensidade de sinal acima ou abaixo de quatro desvios absolutos da mediana (MADs) da intensidade de sinal de todos os *pixels* do ponto. Essa é uma maneira de eliminar artefatos como, por exemplo, partículas de poeira e imperfeições de deposição. Próximo, para cada ponto foi calculada a razão das intensidades de sinal dos canais ($I_{AF555} + I_{AF647}$) e em seguida construído um gráfico que representa a razão de cada ponto em função da intensidade média do sinal nos dois canais [gráfico M vs S, onde $M = \log_2(I_{AF647} / I_{AF555})$ e $S = \frac{1}{2} * \log_2(I_{AF647} + I_{AF555})$], para cada lâmina. Neste espaço, as razões foram normalizadas usando um algoritmo de regressão linear com peso local (LOWESS), implementado previamente pelo Dr. Ricardo Zorzetto N. Vencio (Koide *et al.*, 2004b; Koide *et al.*, 2006a), para corrigir influências sistemáticas nos dados, originadas de pequenas diferenças na marcação e/ou eficiência de detecção entre os corantes fluorescentes (Yang *et al.*, 2002). Após normalização, os dados com valor $S < 3,0$ foram filtrados devido ao baixo sinal, próximo ao ruído. Análises posteriores e seleção de CDS foram feitas com os dados normalizados e filtrados.

Inicialmente, a variação experimental intrínseca das razões dos canais foi estimada a partir de duas hibridizações homotípicas (a mesma amostra é marcada com AF555 e AF647) com cDNA de culturas não tratadas (células em meio PWG, tempo 0h do experimento), onde a razão esperada deve ser próxima a 1 para toda a faixa de intensidade de sinal. Intervalos de confiança de 99,5% foram estimados para faixas de intensidade e os pontos fora dos intervalos de confiança foram considerados promissores, ou diferencialmente expressos. Um ponto dentro de um intervalo de confiança, embora possivelmente apresente razão diferente de zero ($\log_2$ da razão de expressão), foi considerado igualmente expresso nas duas condições. Desta forma foi estimado o ruído das razões de hibridização ao mesmo tempo em que foram preservados os genes que são apenas levemente diferencialmente expressos, uma alternativa considerada mais segura do que o limite "duas vezes mais ou menos expresso" comumente utilizado em estudos com microarranjo (Koide *et al.*, 2004b; Koide *et al.*, 2006a; Zaini, 2007; Zaini *et al.*, 2008).

Uma CDS foi considerada significativamente mais expressa em relação ao controle

(sem tratamento ou tempo zero) se a mediana das réplicas, subtraída de sua MAD (desvio absoluto da mediana), estiver fora dos intervalos de confiança e possuir valor $P < 0,05$ em um teste t contra a hipótese nula $H_0: M \leq$ valor superior do intervalo de confiança para esta intensidade de sinal. Correspondentemente, uma CDS foi considerada significativamente menos expressa em relação ao controle se a mediana das réplicas, subtraída de sua MAD, estiver fora dos intervalos de confiança e possuir valor $P < 0,05$ em um teste t contra a hipótese nula $H_0: M \geq$ valor inferior do intervalo de confiança para esta intensidade de sinal.

3.7. Transcrição reversa seguida de reação em cadeia quantitativa da polimerase em tempo real (RT-qPCR em tempo real)

Os ensaios de RT-qPCR em tempo real para CDS selecionadas foram realizados com iniciadores específicos planejados com o auxílio do programa *Primer Express V2.0*, visando a geração de produtos com tamanho de aproximadamente 150 pares de bases. As sequências dos oligonucleotídeos utilizados estão apresentadas na Tabela 2.

O RNA utilizado como molde para a síntese do cDNA foi obtido como descrito anteriormente a partir de experimentos independentes daqueles destinados às análises de microarranjos (réplicas biológicas). A síntese do cDNA foi realizada utilizando-se 5µg de RNA total, proveniente de amostras incubadas ou não com estreptomicina por 15 e/ou 30 minutos e pré-tratado com 1 U de DNase RQ1 (Promega) para cada 1µg de RNA. Para a síntese do cDNA, acrescentamos 1µL (~500ng) de iniciadores randômicos, 1µL da mistura de dNTP a 10mM e água DEPC para um volume final de 13µL. Essa reação foi incubada a 65°C durante 5 minutos e em seguida, transferida para o gelo por pelo menos 1 minuto. A este volume foram adicionados 4µL de tampão 5x para a síntese de primeira fita do cDNA, 1µL de DTT (Ditiotreitol) a 0,1M, 1µL de RNase OUT (20U) e 1µL (200 U) da enzima transcriptase reversa Superscript III (Invitrogen). A reação foi então incubada a 25°C por 5 minutos e em seguida, a 50°C por 60 minutos. A inativação da reação foi realizada através de incubação a 70°C durante 15 minutos. Em seguida, o cDNA recém sintetizado foi tratado com RNaseH (Fermentas Life Science), diluído com água livre de nuclease para 125 µL e armazenado a -20°C até o uso.

Para cada reação de RT-qPCR, foram utilizados 50 ng de cDNA, 3µL da mistura de iniciadores (direto e reverso) específicos para cada CDS, o que corresponde a uma concentração de 600nM, e 7,5µL do fluoróforo (*SYBR Green* – Applied Biosystems). Para todas as reações foram incluídos controles negativos do RNA utilizado como molde para a síntese de cDNA e dos iniciadores utilizados. As reações de amplificação foram realizadas em um termociclador em tempo real (Applied Biosystems, modelo Gene Amp 5700 ou modelo 7300), utilizando-se o seguinte programa: 2 minutos a 50°C, 2 minutos a 95°C e 50

ciclos de 15 segundos a 95°C e 30 segundos a 60°C. O *software* do equipamento gera curvas de amplificação e dissociação dos iniciadores para cada amostra analisada e oferece o cálculo automático do CT (*Cycle Threshold*) que corresponde ao número de ciclos de reação realizados até que a amplificação seja detectada a partir da emissão de fluorescência. A quantidade de cDNA das amostras foi normalizada de acordo com a expressão do gene controle (*dnaQ*; XF2157) como descrito por (Zaini *et al.*, 2008). A relação entre a expressão do gene de interesse e a expressão do gene controle foi calculada pelo método $2^{-\Delta\Delta C_T}$, conforme proposto por (Livak e Schmittgen, 2001). Os experimentos foram realizados em triplicata e para a maioria das CDS foram considerados experimentos representativos de duas réplicas biológicas.

Tabela 2. Iniciadores utilizados nas reações de RT-qPCR

Identidade da CDS	Iniciadores diretos	Iniciadores reversos
XF0026	TCCTGCACGGCCACAAC	CTGCAATTGGATCATGGATTGA
XF0045	GCCGGTGGCGACTAAGG	CGACGTAGCGGCAATCAAC
XF0107	CAAGCGCGACGGTCGTA	TCAATACCACGCGCGATTTC
XF0121	CGACGCAGGTGCATAAGGA	CAACATCAGCTCGCAAAGCA
XF0169	CTCGCGGCTCAGCACACT	CGCTTCCATCTGCACACAAG
XF0216	CTTTAGAGCCAGCTGCTTTGG	CGGGCAGGACGACCAGTA
XF0274	CCAAAGAATCAATCACGGCAAT	GACGGAGCGCTGGATCAC
XF0287	CGGGCTTCGACAAACTTGA	CAGCGCTCCCTTCTTCCATA
XF0305	TTCATCGTGCCTTGGACTCA	GCCCTTCAGAGCATCGACTTC
XF0340	GTGGAGCATGTTTTGGTTCGT	GCTCGCGCTCTTTACCTTCA
XF0371	GCATCTTGGTCGGCTTTTTTA	ACCAGTGGCTCCGGTAACC
XF0497	CGCCTTCGTCACGGAGATAT	GATCCGGTCCCATAGAAGTTTTT
XF0599	CAGCACACGGACACAAATCAG	GCCGCCCTTTAAGAATTTCTG
XF0615	AACATCCGCGGCATCATC	ATCTTCCAGCATGGCTTTGC
XF0616	CGAAGTTTCCCCTGGTGGTA	GCCTCCCTCAATCTGCAATACC
XF0760	TGAGCCGGTTTACCTTGATGTC	TTCAAGTTCACCGACGGAGAA
XF0767	CCCGGTACGGTTGATGTTG	ACGGCGATGACTTCACCTTT
XF0794	CGGTGGTCTACGCGAATTCT	ATCCATCTGCAACCAACTCACA
XF0800	TGGCGCTTTATAAGGCTTCGT	AGGTCGTCTTGCAGGCTGTACT
XF0889	CTGGATCAAGACAAGGTCAAGGA	CACCCCTTCCTCAGCTACGA
XF1115	GCTTGGAGGAGGCTTCGAA	CACCACCCTTACGAGCAAAAG

XF1171	TGGTTCTGGTTTGGGTAAAACA	TGGAGGCGTTTTAGCTTCGT
XF1407	GGGAGGTATCCGATGCACAT	GCGGTTGATGAGCTGTTCGT
XF1408	CGCTGTTGGAACGTGCATT	AACGAAGCACAGGCAAACG
XF1589	CGGGTGTTGCCGTTCAA	TCGGGTCCTCTATCGTGAGAA
XF1743	CGCCCGATCATCCGTATC	GGTGGAGGTGCCGAACAC
XF1827	CGGGTGGCTCTGGTACCA	TATTTGGATGCGTGGATCGA
XF2048	TTAGCGACGCTGCGACACT	GCACTGGCAAGCTGGTTTG
XF2438	TGCGAGGTGGTGAATATTGG	TGGATAACGCCGCGAAA
XF2466	CGCGCCTGACTCCTGAA	TTGGAGGCCCGTGGATT
XF2628	GGCCCAGGATAAGTTCAAACG	GCCTTAAACTCGCCACCAAA
XF2637	CTCCAGCGTCGATTTTGTTGA	TTGCCACCTTCTCTGTGTTG
XF2753	CGTTTTGCTGCAGCGATTAA	CCACCCCTGATTGAGGATTG
XFa0052	TTGTCGGACGATTTCAATGC	TCGACGCTCAAAGCTTCTCT

3.8. Infecção de *Nicotiana clelandii* com *X. fastidiosa*

As plantas de tabaco (*Nicotiana clelandii*) utilizadas nestes ensaios foram produzidas em recipientes estéreis e mantidas em estufas livres de insetos entre 25 e 30°C no Fundecitrus (Araraquara, São Paulo). Para os testes de inoculação (Lopes *et al.*, 2000) foram utilizadas plantas de tabaco com aproximadamente 2 meses de vida. Estas plantas não receberam água por três dias, antes da inoculação de *X. fastidiosa* para facilitar a captação da suspensão bacteriana pela planta. Foram feitas incisões com agulha de insulina na região dos pecíolos e 15 cm acima desse ponto (Figura 7). Vinte µL de suspensão bacteriana (cepas 9a5c ou J1a12 cultivadas em PWG), de meio PWG apenas, ou de PBS foram adicionados a cada um dos pontos de incisão com o auxílio de uma micropipeta com filtro nas ponteiros. Como controle um grupo de plantas não foi inoculado. Imediatamente após o preparo das suspensões bacterianas para inóculo, uma alíquota foi semeada em meio PWG sólido para avaliação da viabilidade celular.

As plantas foram inspecionadas a cada duas semanas para a verificação de lesões na superfície abaxial das folhas, bem como para os sintomas típicos da CVC em tabaco (Lopes *et al.*, 2000).

Figura 7. Infecção de *Nicotiana clelandii* com *Xylella fastidiosa*. Plantas de tabaco foram inoculadas com as cepas 9a5c e /ou J1a12. Vinte μL de suspensão bacteriana foram aplicados no caule da planta em dois pontos diferentes (total de 40 μL por planta) **(A)**. A absorção foi garantida pela cuidadosa perfuração desses locais com uma agulha de insulina **(B)**. Foram inoculadas plantas de 2 meses de idade e aproximadamente 20 cm de altura **(C)**.

3.9. Pirosequenciamento do genoma de *X. fastidiosa*.

3.9.1. Purificação de DNA genômico de *Xylella fastidiosa*.

As cepas de *X. fastidiosa*, 9a5c e J1a12, foram cultivadas em meio complexo PWG como descrito no item 3.1. Para purificação de DNA genômico, as células foram coletadas por centrifugação a 3.000 xg por 5 minutos a 4°C, a partir de 250 mL de cultura em meio PWG. O precipitado de células foi lavado com 5 mL de TAS (Tris-HCl 50mM pH 8,0, EDTA 0,5mM, NaCl 150mM) por 2 min a 12000 rpm. O sobrenadante foi descartado e o precipitado de células, suspenso com 450µL de TAS, seguindo-se adição de 4µL de proteinase K (20mg/mL). Após incubação a 30°C por 60 minutos, foi adicionado 50µL de SDS 1% e a suspensão foi misturada por inversão do tubo. Após incubação a 50°C por 60 min, foi adicionado igual volume de clorofórmio/álcool isoamílico 24:1 (v/v), seguindo-se agitação vigorosa e separação da fase aquosa por centrifugação a 10.000 xg por 10 minutos. A fase aquosa foi transferida para um novo tubo e extraída com igual volume de fenol/clorofórmio/álcool isoamílico 25:24:1 (v/v). A fase aquosa foi separada por centrifugação nas mesmas condições e o DNA foi precipitado pela adição de 0,1 volume de acetato de potássio 3 M, pH 5,5 e 0,6 volume de isopropanol à fase aquosa. O tubo foi centrifugado a 5.000 xg por 20 min a 4°C e o precipitado foi lavado em 2 mL de etanol 70% (v/v) resfriado a -20°C e centrifugado nas mesmas condições. O etanol foi desprezado, o precipitado foi seco por centrifugação a vácuo (Eppendorf concentrator 5301) por 5 minutos e ressuspenso em 300 µL Tris-HCl 10 mM, pH 8,0. Para garantir maior pureza da amostra, foi realizada diálise (48 a 72h) em TE pH 8,0, a 4°C, trocando-se o tampão a cada 24h.

Alternativamente, foi utilizado *kit* comercial de isolamento de DNA genômico (*Wizard Genomic DNA Purification kit*, Promega), seguindo-se o protocolo do fabricante.

3.9.2. Análise do DNA genômico das cepas 9a5c e J1a12.

Um pré-requisito para o bom resultado do pirosequenciamento é que a qualidade e a quantidade de DNA sejam seguramente determinadas. A qualidade e quantidade do DNA genômico purificado foram avaliadas pela medida das absorbâncias a 230nm, 260nm e 280nm e também pela análise eletroforética em gel de agarose e no aparelho BioAnalyser (Agilent) como descrito no item 3.4.

Para eletroforese em gel de agarose, aproximadamente 200ng de DNA foram misturados ao tampão de amostra [corante azul de bromofenol 0,25% (p/v) e sacarose 40% (p/v)] com volume final de 10 µL. A eletroforese foi realizada a 100 Volts em gel de agarose (0,8% a 1,2%) contendo brometo de etídio (10mg/mL), utilizando-se como tampão de corrida

TAE 1X (Ausubel e Frederick, 1995). Ao final da corrida, o gel foi fotografado em um transiluminador com luz UV e o tamanho dos fragmentos foi estimado com base na comparação à mobilidade de fragmentos de DNAs com tamanhos já conhecidos (1Kb Ladder - Invitrogen) (Figura 8).

Figura 8. Análise eletroforética de amostra de DNA genômico de *X. fastidiosa*. A foto ilustra a qualidade da amostra de DNA genômico utilizada para o pirosequenciamento, separada em gel de agarose 1%. Os marcadores de tamanho (DNA 1kb Ladder) estão na canaleta 1 e a canaleta 2 corresponde à amostra de DNA genômico de cepa J1a12 de com 15 passagens em cultivo.

Antes de iniciarmos a construção das bibliotecas de DNA genômico de *Xylella fastidiosa*, as amostras de DNA a serem utilizadas foram também quantificadas com *Quant-iT PicoGreen* (Invitrogen). Para isso, o reagente fluorescente *PicoGreen* 200x foi diluído com TE 1x para a concentração de 1x e foram preparadas diluições de uma amostra de DNA de fago lambda (100µg/mL) utilizada como padrão pelo *Quant-iT PicoGreen dsDNA Kit* (Invitrogen) para a construção de uma curva padrão (Tabela 3). Os ensaios para quantificação dos DNAs genômicos provenientes de culturas de 9a5c com 14 passagens e de J1a12 com 15 e 24 passagens foram realizados separadamente de acordo com as indicações do fabricante. O *PicoGreen* é excitado no comprimento de onda de 480nm e emite fluorescência no comprimento de onda de 520nm. A quantificação de DNA por este método é mais precisa que aquela feita em espectrofotômetro a 260nm, pois este fluoróforo só é incorporado pelo DNA dupla-fita. A quantificação por *PicoGreen* foi realizada em fluorímetro e a determinação da concentração de DNA de cada uma das amostras foi feita a partir da construção de uma curva padrão de concentração. Os valores de unidade de

absorbância obtidos para cada amostra foram multiplicados pelo valor do coeficiente angular dessa reta (Figura 9 e Tabela 4).

Tabela 3. Preparo da curva padrão *high-range* para quantificação de DNA genômico com o Quan-iT PicoGreen (Invitrogen)

Volume de TE 1x (μL)	Volume de DNA padrão (2μg/mL) (μL)	Volume de Quant-iT PicoGreen 1x (μL)	Concentração Final do padrão Quant-iT PicoGreen
0	1000	1000	1μg/mL
900	100	1000	100ng/mL
990	10	1000	10ng/mL
999	1	1000	1ng/mL
1000	0	1000	Branco

Figura 9. Quantificação das amostras de DNA genômico de *X. fastidiosa* com PicoGreen (Invitrogen). O eixo y representa a intensidade de fluorescência do sinal emitido em cada diluição das mostras que constituem a curva padrão de dosagem de DNA. O eixo x representa a concentração de DNA padrão do *Quant-iT PicoGreen dsDNA Kit* (Invitrogen) nas diluições que compõem a curva padrão *high range*.

Tabela 4. Determinação da concentração das amostras de DNA genômico de *X. fastidiosa*.

Biblioteca	Sinal de Fluorescência	Concentração ng/ μ L na solução de dosagem	Fator de diluição da solução de dosagem	Concentração ng/ μ L	Volume total	Massa total	Massa nebulizada
J1a12 15P	160	20,64	2	41,27	1500 μ L	61905ng	4127,27ng
J1a12 24P	347,7	26,77	2	53,55	200 μ L	10710ng	5354,55ng
9a5c 14P	133,3	69,43	2	138,86	1500 μ L	208290ng	Diluição 30x = 4628ng

3.9.3. Genotipagem de *Xylella fastidiosa*.

A genotipagem das cepas de *X. fastidiosa* visando a confirmar sua autenticidade, foi realizada através de ensaios de PCR que detectam diferenças genômicas previamente detectadas entre as cepas 9a5c e J1a12 (Koide *et al.*, 2004b). Utilizamos iniciadores para as regiões externas às CDS XF0077, que codifica um precursor de adesina de fímbria, e XF1250, que codifica uma arginina desaminase, as quais estão ausentes no genoma de J1a12 (Koide *et al.*, 2004b) (Tabela 5).

Para a amplificação dessas CDS, 100ng DNA genômico das amostras de DNA de 9a5c com 14 passagens e J1a12 com 15 e 24 passagens foram utilizados como molde em ensaios de PCR, com 1 μ L de solução de dNTPs 10mM, 1,5 μ L de MgCl₂ 25mM, 10 μ L de tampão 2x da *Taq* DNA Polimerase (Invitrogen), 1 μ L de cada iniciador 10pmol/ μ L e água MilliQ estéril filtrada para o volume final da reação de 50 μ L. As reações foram submetidas à desnaturação a 95°C, 5 min e a 35 ciclos de 95°C, 45 seg; 60°C, 30 seg; 72°C, 2 min. Ao final destes ciclos, as reações foram incubadas a 72°C, 10 min para garantir a extensão de todos os *amplicons* formados e então, a 4°C até sua análise por eletroforese em gel de agarose 1%.

A figura 10 mostra a eletroforese em gel de agarose 1% dos *amplicons* obtidos após a PCR do DNA das diferentes cepas. Como esperado, a amplificação com XF0077F e XF0077R gerou *amplicons* de 1599pb em 9a5c e 549pb em J1a12 e a amplificação com XF1250F e XF1250R gerou *amplicons* de 1122pb em 9a5c e 159pb em J1a12.

Tabela 5. PCR para genotipagem de 9a5c e J1a12

	Iniciadores	Tamanho do amplicon em 9a5c	Tamanho do amplicon em J1a12
$\Delta 77F$	ACTAAGCGCGACAATGAGGA	1599pb	549pb
$\Delta 77R$	GTGGTGTTTGGGCAAGCTAT		
$\Delta 1250F$	AATGGCAAGGTGGAAATCAG	1122pb	159pb
$\Delta 1250R$	CAACGCACATCCCCTAAGTC

Figura 10. Genotipagem das cepas 9a5c e J1a12. Os produtos da reação de PCR com os iniciadores XF0077 (F e R) e XF1250 (F e R) do DNA de J1a12 e 9a5c estão mostrados nas canaletas canaleta 2 e canaletas 4 a 6, respectivamente. A canaleta 3 ao controle (PCR sem adição de DNA). Os marcadores de tamanho (1Kb Ladder, Invitrogen) foram aplicados na canaleta 1. A eletroforese foi realizada em gel de agarose 1% contendo brometo de etídio. As setas largas apontam os amplicons de 1599 pb e 1122 pb e as setas finas apontam os amplicons de 549 pb e 159 pb (de cima para baixo).

3.9.4. Pirosequenciamento de DNA

A tecnologia de pirosequenciamento de DNA 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). O pirofosfato liberado a cada reação de incorporação pela DNA polimerase é utilizado na síntese de ATP pela ATP sulfúrilase. 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 seqüenciamento (Margulies *et al.*, 2005; Ahmadian *et al.*, 2006). O equipamento 454 *Genome Sequencer FLX* (454 Life Sciences/Roche), permite, atualmente, a obtenção de sequências (*reads*) de 400-500 nucleotídeos, sendo possível o sequenciamento de até 400-600 milhões de bases de alta qualidade por corrida com a utilização dos reagentes da série titânio (*GS FLX Titanium Series Reagents*) (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 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, separando-se o óleo. As microesferas são depositadas e imobilizadas em uma placa (*PicoTiterPlate*) contendo 1,6 milhões de poços; cada poço tem a dimensão exata para comportar uma única microesfera. A placa é em seguida acoplada ao sequenciador para as reações de pirosequenciamento. Após ~8-10 horas de corrida, 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 ~8-10 horas adicionais de análise computacional, as imagens são convertidas em sequências, cada uma derivada das imagens de um poço da placa. Em

média são obtidas aproximadamente 1,2 milhão de sequências, cada uma com mais de 400 bases com 99% de confiança (*Phred score* > 20).

As tecnologias de sequenciamento ultra-rápidas e com alto desempenho (454 Life Sciences/Roche, Illumina/Solexa e SoLiD) têm sido utilizadas com sucesso no sequenciamento de genomas e metagenomas e também de transcritomas completos. Em pouco mais de três anos, a literatura acumulou centenas de artigos descrevendo resultados obtidos a partir da aplicação destas tecnologias. Destes vários descrevem o pirosequenciamento completo de genomas bacterianos de cepas de *Mycoplasma genitalium*, *Salmonella Typhi*, *Staphylococcus aureus*, *Sinorhizobium meliloti*, *Campylobacter jejuni* (Margulies *et al.*, 2005; Highlander *et al.*, 2007; Poly *et al.*, 2007; Holt *et al.*, 2008; Stiens *et al.*, 2008), possibilitando a análise da extensa variação genômica entre isolados destas espécies bacterianas.

O protocolo de pirosequenciamento e resultados obtidos em cada uma das etapas deste procedimento estão apresentados como material suplementar (Anexos).

3.9.10. Análises bioinformáticas

Para pesquisa de identidade/similaridade de sequências de DNA ou de proteínas utilizamos os programas BLASTn, BLASTp e BLASTx (*Basic Local Alignment Search Tool*) (Altschul *et al.*, 1990; Altschul *et al.*, 1997) 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* Genome Comparative Project (<http://www.xylella.lncc.br/>).

Para o alinhamento de sequências de aminoácidos ou de nucleotídeos utilizamos o programa ClustalX ou ClustalW (Thompson *et al.*, 1994; Thompson *et al.*, 1997). O planejamento de alguns oligonucleotídeos foi realizado com o programa PRIMER3 (Rozen e Skaletsky, 2000).

A montagem de *contigs* a partir das sequências obtidas no sequenciador 454 foi realizada utilizando-se o Programa *Newbler* (454 Life Sciences/Roche) ou utilizando-se SABIA (*System for Automated Bacterial (genome) Integrated Annotation*) (Almeida *et al.*, 2004), uma plataforma desenvolvida no Laboratório de Bioinformática do LNCC (Laboratório Nacional de Computação Científica, <http://www.labinfo.lncc.br/>). Esta plataforma integra uma série de programas de montagem entre eles o *Newbler*. Como referência para a montagem, foi utilizada a sequência genômica da cepa 9a5c (Simpson *et al.*, 2000), utilizando-se o módulo *comparative* da plataforma SABIA e/ou *GS Reference Mapper* (454 Life Sciences/Roche).

Análises comparativas dos genomas de *X. fastidiosa* foram realizadas utilizando-se o servidor RAST (*Rapid Annotation using Subsystem Technology*) (Aziz *et al.*, 2008).

4. Resultados e Discussão

4.1. Análise do perfil de expressão gênica global de *Xylella fastidiosa* em resposta ao tratamento com dose subletal de estreptomicina

Em concentrações subinibitórias, agentes antimicrobianos modulam a expressão gênica bacteriana, sendo que o conjunto de genes que é modulado depende tanto da cepa bacteriana e como da natureza do agente antimicrobiano (Goh *et al.*, 2002; Bader *et al.*, 2003; Hong *et al.*, 2003; Ng *et al.*, 2003; Shaw *et al.*, 2003; Hoffman *et al.*, 2005; Linares *et al.*, 2006; Jones *et al.*, 2009). Além disso, alterações no perfil de expressão gênica em resposta à exposição a concentrações subletais de antibióticos correlacionam-se com alterações fenotípicas. Por exemplo, a exposição de *Pseudomonas aeruginosa* a doses muito baixas do peptídeo antimicrobiano LL-37 regula negativamente a expressão de genes associados a *quorum sensing* e previne a formação de biofilme (Overhage *et al.*, 2008). Por outro lado, a exposição de *P. aeruginosa* a doses subletais de tetraciclina, tobramicina ou norfloxacin resultou no aumento dos níveis de transcritos de genes associados colonização e virulência e na formação de biofilme (Hoffman *et al.*, 2005; Linares *et al.*, 2006).

Dados obtidos em nosso laboratório mostraram que a exposição de *Xylella fastidiosa* a concentrações subletais de peptídeo antimicrobiano gomesina induz mudanças no perfil de expressão gênica global, em particular um aumento nos níveis de transcritos de genes relacionados a produção de biofilme e virulência de *Xylella fastidiosa*, cepa 9a5c (Fogaça *et al.*, 2010). Paralelamente, aos estudos realizados com a gomesina, nós investigamos quais os efeitos da exposição de *X. fastidiosa* a doses subletais de estreptomicina, um aminoglicosídeo utilizado na agricultura durante décadas para o controle de diversos patógenos, tais como *Erwinia amylovora*, *Pseudomonas syringae* e *Xanthomonas vesicatoria* (McManus *et al.*, 2002; Vidaver, 2002).

A dose subletal de estreptomicina utilizada em nossos experimentos foi definida após ensaios para determinação da suscetibilidade de duas cepas de *X. fastidiosa* (cepas 9a5c e J1a12) a este e outros antibióticos, os quais foram realizados por A.C. Fogaça em nosso laboratório (Fogaça *et al.*, 2010). Como mostra a Tabela 6, a concentração bactericida mínima (MBC) para estreptomicina está na faixa de 3,1 a 6,2 µg/mL, após tratamento por 18 h, sendo que não observamos diferenças entre as duas cepas testadas. Desta forma, escolhemos como dose subletal 1 µg/mL, uma concentração 3 a 6 vezes menor que a MBC verificada para estreptomicina. Ensaio adicionais confirmaram que a exposição de *X. fastidiosa* a 1 µg/mL por tempos que variaram de 15 min a 18 horas não afeta sua viabilidade enquanto que concentrações maiores (5 e 10 µg/ml) tem efeito oposto (Figura 11).

Tabela 6. Perfil de suscetibilidade da *X. fastidiosa* a diferentes antibióticos

Antibiótico	Concentração bactericida mínima (MBC) ^a contra <i>X. fastidiosa</i> , cepas 9a5c e J1a12	
	µg/mL	µM
Gomesina	4,5 – 9	200-400
Ampicilina	$1,25 \times 10^3$ – $2,5 \times 10^3$	$3,35 \times 10^3$ – $6,7 \times 10^3$
Estreptomicina	3,1 – 6,2	6,4 – 12,8
Tetraciclina	7,8 – 15,6	10,7 – 21,4

^aA concentração bactericida mínima para os antibióticos está expressa como a faixa *a-b*, onde *a* corresponde a maior concentração onde foi ainda foi verificado crescimento bacteriano e *b* corresponde a menor concentração que matou 100% das células. (Tabela extraída de (Fogaça *et al.*, 2010).

Figura 11. Ensaio de viabilidade celular de *X. fastidiosa* exposta a doses subletais de estreptomicina. Células de *Xylella fastidiosa* crescidas em meio PWG foram tratadas com 0,1, 1, 5 e 10 µg/mL (A, B, C e D respectivamente) de estreptomicina pelos tempos indicados (15 e 60 minutos, 3 e 18 horas) e imediatamente plaqueadas em meio PWG-água. Como controle, amostras de células não tratadas (adição de H₂O MilliQ) foram coletadas e plaqueadas ao final dos tempos indicados. O crescimento foi avaliado após sete dias de incubação a 28° C.

O perfil transcricional de *X. fastidiosa*, cepa 9a5c, em resposta ao tratamento com dose subletal de estreptomicina foi avaliado utilizando-se microarranjos de DNA em lâminas de vidro representando 2608 sequências codificadoras (CDS), isto é 91,6%, das 2848 CDS previamente anotadas no genoma da cepa 9a5c conforme descrito (Simpson *et al.*, 2000; Koide *et al.*, 2004b). Células de *X. fastidiosa* cepa 9a5c, com no máximo 20 passagens de cultivo em meio PWG, foram incubadas com 1µg/mL de estreptomicina por 15, 30, e 60 min. O RNA total de células tratadas ou não com estreptomicina nos tempos indicados foi isolado e utilizado para síntese do cDNA fluorescente utilizado na hibridização dos microarranjos. Os dados obtidos correspondentes a pelo menos duas réplicas biológicas com duas réplicas técnicas cada, estão sumarizados na Tabela 7.

Nas diversas hibridizações que realizamos ~65-70% de CDS do total de CDS representadas no microarranjo foi detectado em todos os tempos amostrados, isto é com sinal fluorescente acima do *background* (Tabela 7). Destas, 136 CDS foram consideradas diferencialmente expressas após o tratamento com 1µg/mL de estreptomicina, segundo os critérios de análise detalhados no item 3.6. (Procedimentos Experimentais) Dentre as CDS diferencialmente expressas, 27 foram positivamente moduladas e 109 foram negativamente moduladas ao longo do tratamento por 15, 30 e 60 min com 1µg/mL estreptomicina (Tabelas 8 e 9). Como já mencionado, a exposição de *X. fastidiosa* este tratamento não afeta a sua viabilidade celular. A porcentagem de CDS moduladas em relação a CDS detectadas que observamos em nossos experimentos foi de ~8%, sendo que observamos modulação negativa para 80% (109/136) das CDS diferencialmente expressas. A exposição de *Yersinia pestis* a doses mais altas de estreptomicina (80µg/mL, 10 vezes maior que a MBC determinada para esta bactéria) resultou na modulação de ~14,5% das CDS analisadas através de microarranjos de DNA (Qiu *et al.*, 2005). Estes autores reportam que, das CDS moduladas após o tratamento de *Y. pestis* por 30 min, 58% (201/345) foram reprimidas. A exposição de *Streptococcus pneumoniae* a doses subletais de puromicina, tetraciclina, cloranfenicol ou eritromicina resultou, em média, na modulação da expressão de 10% das CDS analisadas, sendo que mais da metade (55-65%) destas foram reprimidas (Ng *et al.*, 2003).

Tabela 7. Total de CDS de *X. fastidiosa* expressas e diferencialmente expressas em resposta ao tratamento com 1µg/mL de estreptomicina.

CDS representadas no microarranjo	2608	
CDS expressas ^a	0 min	1734
	15 min	21
CDS positivamente moduladas	30 min	7
	60 min	-
	15 min	67
CDS negativamente moduladas	30 min	43
	60 min	8
Total of CDS positivamente moduladas ^b	27	
Total of CDS negativamente moduladas ^b	109	

^a O número de CDS consideradas expressas corresponde ao número de CDS que apresentaram sinal de fluorescência acima *background*. O RNA total utilizado na hibridização homotípica (controle) foi isolado de células coletadas antes da adição de estreptomicina (tempo 0 min). O número de CDS detectadas não apresentou alteração relevante ao longo dos pontos amostrados

^b Total de CDS únicas

Tabela 8. CDS moduladas positivamente em resposta ao tratamento com 1µg/mL de estreptomicina

Identidade da CDS ^a	Nome do gene ^b	Produto predito ^b	Razão de Expressão ^c	M ^c	Tempo (min) ^d
XF0026	<i>aroG</i>	phospho-2-dehydro-3-deoxyheptonate aldolase	2,03	1,02	15
XF0045	<i>mfd</i>	transcription-repair coupling factor (ATP-dependent helicase Mfd)	3,01	1,59	15
XF0107	<i>rpsP</i>	30S ribosomal protein S16	2,49	1,32	15
<u>XF0195</u>	-	hypothetical protein	2,90	1,54	15
XF0201	-	SAM-dependent methyltransferase family protein	2,44	1,29	15
XF0315	<i>nuoK</i>	NADH-quinone oxidoreductase chain K	1,98/10,27	0,75/ 3,36	15/30
XF0461	-	conserved hypothetical protein	2,73	1,45	15
XF0599	<i>fiu</i>	catecholate siderophore receptor Fiu (TonB-dependent receptor Fiu)	2,54	1,34	15
XF0615	<i>groL</i>	chaperonin GroEL (60kDa chaperonin)	1,55	0,63	15
XF0616	<i>groS</i>	co-chaperonin GroES (10 kDa chaperonin)	1,91	0,93	15
<u>XF0642</u>	-	hypothetical protein	2,30	1,20	30
XF0647	<i>nth</i>	endonuclease III	3,73	1,90	30
XF0759	<i>amiC</i>	N-acetylmuramoyl-L-alanine amidase	2,75	1,46	30
XF0794	<i>murF</i>	UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate—D-alanyl-D-alanyl ligase	2,16	1,11	30
XF1743	<i>est</i>	esterase; giCVC	3,63	1,86	15
XF1827	<i>ohr</i>	organic hydroperoxide resistance protein	1,75	0,81	15
XF1833	<i>ubiB</i>	ubiquinone biosynthesis protein	4,50	2,17	15
XF2048	<i>trbJ2</i>	conjugal transfer protein TrbJ	1,62	0,70	15
XF2438	<i>rpsA</i>	30S ribosomal protein S1	2,68	1,42	15
XF2466	<i>pglA</i>	polygalacturonase (PGA) (Pectinase)	2,22	1,15	30
XF2628	<i>tufA</i>	elongation factor Tu	1,75	0,81	15
XF2629	<i>fusA</i>	elongation factor G (EF-G)	2,07	1,05	15
XF2637	<i>rplK</i>	50S ribosomal protein L11	2,07	1,05	30
XF2706		peptidase M28 family protein	2,37	1,24	15
XF2724	<i>hsdM1</i>	type I restriction-modification system	2,39	1,26	15
Xfa0025		acetyltransferase GNAT family protein	3,01	1,59	15
Xfa0052	<i>vapD</i>	virulence-associated protein D	2,73	1,45	15

^a Identidade da CDS segundo (Simpson *et al.*, 2000). Estão sublinhadas as CDS eliminadas da re-anotação do genoma da *X.fastidiosa* cepa 9a5c disponível em <http://www.xylella.Incc.br>

^b Nome do gene e o produto predito de acordo com (Simpson *et al.*, 2000) considerando-se, para algumas CDS, a recente re-anotação do genoma da cepa 9a5c.

^c Razão de expressão em relação a células não tratadas. M= Log₂ da razão de expressão.

^d Tempo após tratamento com estreptomicina em que foi observada a modulação da expressão.

Tabela 9. CDS moduladas negativamente em resposta ao tratamento com 1µg/mL de estreptomicina

Identidade da CDS ^a	Nome do gene ^b	Produto predito ^b	Razão de Expressão ^c	M ^c	Tempo (min) ^d
XF0012	<i>exbD2</i>	biopolymer transport ExbD2 protein	0,53	-0,91	60
XF0068	<i>ubiA</i>	4-hydroxybenzoate octaprenyltransferase	0,25	-2,02	30
XF0073	<i>ffh</i>	signal recognition particle protein	0,63	-0,68	15
XF0083	<i>fimA</i>	fimbrial subunit precursor	0,33	-1,58	15
XF0104	<i>htrB1</i>	lipid A biosynthesis lauroyl acyltransferase	0,19	-2,40	30
XF0112	-	hypothetical protein	0,25/0,25	-2,01-2,00	15/30
XF0121	-	2-nitropropane dioxygenase (NPD) family protein	0,36/0,40	-1,48/-1,31	15/30
XF0145	<i>sdh</i>	Serine 3-dehydrogenase	0,29	-1,76	30
XF0246	-	hypothetical protein	0,27	-1,90	30
XF0274	<i>pfkA</i>	6-phosphofructokinase	0,40	-1,31	15
XF0286	<i>sulA</i>	putative sugar nucleotide epimerase	0,10	-3,35	30
XF0298	-	hypothetical protein ^c	0,31/0,25	-1,70/-2,01	15/60
XF0307	<i>nuoC</i>	NADH-quinone oxidoreductase chain C	0,51	-0,96	15
XF0320	<i>citM</i>	Mg ⁺⁺ /citrate complex transporter	0,58	-0,79	15
XF0321	<i>oprO1</i>	porin O precursor	0,37	-1,45	15
XF0323	<i>tctE</i>	two-component system, sensor protein	0,32/0,43	-1,65/-1,23	15/30
XF0328	-	DNA injection protein [Xylella phage Xfas53]	0,21	-2,25	30
XF0339	-	TonB-dependent receptor	0,55	-0,86	15
XF0340	<i>dsbB</i>	disulfide bond formation protein B	0,42	-1,26	15
XF0345	-	hypothetical protein	0,36	-1,46	30
XF0357	-	Xanthomonadales conserved hypothetical	0,48	-1,05	15
XF0359	-	hypothetical protein	0,41	-1,29	15
XF0409	-	hypothetical protein	0,30	-1,72	60
XF0411	<i>nrtB</i>	ABC transporter nitrate permease	0,53	-0,93	15
XF0413	<i>rmuC</i>	DNA recombination protein RmuC homolog	0,19	-2,38	60
XF0474	<i>pilV2</i>	Tfp pilus assembly protein PilV	0,12	-3,09	30
XF0475	<i>pilW2</i>	Tfp pilus assembly protein PilW	0,12	-3,07	30
XF0480	<i>int</i>	phage-related integrase (fragment)	0,34	-1,55	15
XF0483	-	phage-related protein; xfp6	0,30	-1,75	30
XF0485	-	phage-related protein; xfp6	0,48	-1,07	15
XF0488	-	hypothetical protein; xfp6	0,47	-1,08	15
XF0497	-	hypothetical protein; xfp6	0,31	-1,67	30
XF0584	-	conserved hypothetical protein	0,40	-1,32	30
XF0632	-	conserved hypothetical protein	0,56	-0,85	30
XF0766	-	conserved hypothetical protein	0,40	-1,32	15
XF0774	<i>pig7</i>	Xanthomonadin biosynthesis dehydratase protein	0,47	-1,08	15
XF0807	-	hypothetical protein	0,53	-0,91	30
XF0822	<i>gltX</i>	glutamyl-tRNA synthetase	0,46	-1,12	15
XF0838	<i>surA</i>	peptidyl-prolyl cis-trans isomerase	0,67	-0,58	15
XF0852	<i>alr</i>	alanine racemase	0,15	-2,72	30
XF0870	-	hypothetical protein	0,24	-2,08	30

XF0881	<i>vanY</i>	D-alanyl-D-alanine carboxypeptidase (peptidase family M15B)	0,23	-2,12	30
XF0885	-	Glycosyl transferase group 1 family protein	0,44	-1,18	15
XF0890	-	hypothetical protein	0,18	-2,51	30
XF0902	-	Xanthomonadales conserved hypothetical	0,21	-2,28	30
XF0945	-	hypothetical protein	0,29	-1,79	30
<u>XF1028</u>	-	hypothetical protein	0,50	-0,99	15
XF1037	<i>ahcY</i>	S-adenosyl-L-homocysteine hydrolase	0,25	-2,01	60
XF1061	<i>eda</i>	2-dehydro-3-deoxy-phosphogluconate aldolase (KDPG-aldolase)	0,63	-0,68	15
XF1090	-	hypothetical protein	0,48	-1,05	30
XF1116	<i>lysAC</i>	Bifunctional diaminopimelate decarboxylase/aspartate kinase	0,56	-0,84	15
<u>XF1122</u>	-	hypothetical protein	0,33	-1,62	30
<u>XF1129</u>	-	hypothetical protein	0,63	-0,68	30
XF1137	<i>nonF</i>	ThiJ/Pfpl family protein	0,31	-1,67	30
XF1140	<i>glmU</i>	Bifunctional protein GlmU: UDP-N-acetylglucosamine pyrophosphorylase / Glucosamine-1-phosphate N-acetyltransferase	0,10	-3,27	30
XF1143	<i>atpD</i>	ATP synthase beta subunit	0,54	-0,89	15
XF1188	<i>clpX</i>	ATP-dependent Clp protease ATP binding subunit Clpx	0,48	-1,05	15
XF1226	-	Xanthomonadales conserved hypothetical	0,56	-0,84	15
XF1243	-	conserved hypothetical protein	0,28	-1,82	15
XF1244	-	hypothetical protein	0,66	-0,59	15
<u>XF1246</u>		hypothetical protein	0,44	-1,17	15
XF1275	<i>phaF</i>	poly(hydroxyalkanoate) granule associated protein	0,51	-0,98	15
XF1374	<i>trpF</i>	N-(5'-phosphoribosyl)anthranilate isomerase	0,29	-1,76	30
XF1408	<i>rpoN</i>	RNA polymerase factor sigma-54	0,35	-1,51	15
XF1506	-	HAM1 protein (putative purine NTP pyrophosphatase, rdgB/HAM1 family)	0,25/0,27	-2,00/1,89	15/30
XF1516	<i>xadA1/uspA1</i>	Surface protein adhesin YadA-like (Xanthomonas adhesin XadA-like protein)	0,39	-1,37	60
XF1589	<i>y4JK</i>	plasmid stabilization protein ; xfp3	0,26	-1,94	15
XF1604	<i>btuE</i>	ABC transporter vitamin B12 uptake permease	0,58	-0,78	15
XF1620	-	Xanthomonadales conserved hypothetical	0,51	-0,99	15
XF1629	-	hypothetical protein	0,75	-0,42	15
XF1661	-	hypothetical protein; xfp4	0,07	-3,76	30
XF1737	-	NADP oxidoreductase, coenzyme F420-dependent family protein	0,15	-2,71	30
XF1738	<i>tlcA</i>	ATP/ADP translocase	0,60	-0,73	15
XF1760	-	phage-related protein	0,51	-0,98	15
XF1806	<i>dinD</i>	DNA-damage-inducible protein	0,37	-1,42	15
XF1812	-	Transglutaminase-like protein	0,31	-1,68	15
XF1835	-	putative reductase	0,23	-2,15	30
XF1855	<i>fum</i>	fumarate hydratase class I	0,28	-1,83	15
XF1884	-	hypothetical protein	0,62	-0,70	15
XF1887	-	conserved hypothetical protein	0,20	-2,30	15

XF1911	-	Xanthomonadales conserved hypothetical	0,53	-0,92	15
XF2012	-	conserved hypothetical protein	0,65	-0,61	15
XF2017	-	conserved hypothetical protein	0,61	-0,70	15
XF2022	<i>sbcB</i>	exodeoxyribonuclease I	0,18	-2,51	30
XF2049	<i>trbI2</i>	conjugal transfer protein TrbI	0,10	-3,35	30
XF2050	<i>trbH2</i>	conjugal transfer protein TrbH	0,41	-1,29	30
XF2144	<i>pstB</i>	phosphate import ATP-binding protein	0,61	-0,71	15
XF2245	-	Xanthomonadales conserved hypothetical	0,04	-4,56	30
XF2318	-	hypothetical protein	0,54	-0,89	15
XF2366	<i>gumE</i>	xanthan biosynthesis polysaccharide polymerase	0,57	-0,80	15
XF2370	<i>gumB</i>	Xanthan biosynthesis polysaccharide transport protein GumB	0,36	-1,48	30
XF2430	<i>guaB</i>	inosine-5'-monophosphate dehydrogenase (IMPDH)	0,67	-0,58	15
XF2434	<i>wbpL</i>	lipopolysaccharide core biosynthesis protein	0,60	-0,73	15
XF2448	<i>ugpB</i>	ABC transporter sugar-binding periplasmic protein	0,30	-1,75	30
XF2463	<i>ccmH2</i>	C-type cytochrome biogenesis protein CcmH	0,45	-1,14	15
XF2472	-	Amidohydrolase family protein	0,55	-0,87	15
XF2477	<i>bioD</i>	dithiobiotin synthetase	0,19/0,24	-2,40/2,06	15/30
XF2514		hypothetical protein ; xfp2	0,52	-0,95	15
XF2547	<i>sucC</i>	succinyl-CoA synthetase, beta subunit	0,43	-1,22	30
XF2548	<i>sucD</i>	Succinyl-CoA ligase [ADP-forming] subunit alpha	0,31	-1,70	30
XF2552	<i>gyrA</i>	DNA gyrase subunit A	0,32/0,37	-1,65/-1,45	15/60
XF2687	-	Xanthomonadales conserved hypothetical	0,48	-1,06	15
XF2688	-	Xanthomonadales conserved hypothetical	0,47	-1,09	15
XF2703	-	hypothetical protein	0,63	-0,67	15
XF2753	-	hypothetical protein	0,39/0,33	-1,36/-1,59	15/30
XFa0006	<i>virB3</i>	conjugal transfer protein VirB3	0,70	-0,52	15
XFa0009	-	pKM101 eex-like protein	0,55	-0,86	15
<u>XFa0028</u>	-	hypothetical protein	0,57/0,55	-0,80/-0,87	15/60
XFa0033	-	Resolvase, N-terminal domain	0,30	-1,74	15

^a Identidade da CDS segundo (Simpson *et al.*, 2000). Estão sublinhadas as CDS eliminadas da re-anotação do genoma da *X. fastidiosa* cepa 9a5c disponível em <http://www.xylella.lncc.br>

^b Nome do gene e o produto predito de acordo com (Simpson *et al.*, 2000) considerando-se, para algumas CDS, a recente re-anotação do genoma da cepa 9a5c.

^c Razão de expressão em relação a células não tratadas. M= Log₂ da razão de expressão.

^d Tempo após tratamento com estreptomicina em que foi observada a modulação da expressão.

O nível dos transcritos de 21 CDS selecionadas a partir dos dados de microarranjos foi também avaliado através de RT-qPCR (Figura 12). Para estes ensaios novas preparações de RNA total foram obtidas de células tratadas ou não com 1µg/mL estreptomicina por 15 e/ou 30 min, uma vez que foram estes os períodos de incubação em que observamos maior variação no perfil de expressão gênica de *X. fastidiosa* (Tabela 7). A Tabela 10 mostra que o perfil transcricional obtido pelas análises de microarranjos foi validado para 16 CDS com os ensaios de RT-qPCR, o que corresponde a uma taxa de validação de ~75%, um valor considerado aceitável em estudos comparativos entre estas plataformas (Petersen *et al.*, 2005; Twellmeyer *et al.*, 2007; Zaini *et al.*, 2008). Observamos uma discrepância na magnitude do valor de razão de expressão para algumas CDS quando comparamos dados de microarranjos e RT-qPCR (CDS XF1743 e XF2048, Tabela 10). Este fato também já foi observado em outros estudos, sendo que, frequentemente, dados de RT-qPCR apresentam maior amplitude do que dados de microarranjos (Petersen *et al.*, 2005; Twellmeyer *et al.*, 2007; Zhang *et al.*, 2008). De qualquer modo, embora as magnitudes dos dados de microarranjos sejam menores, esta metodologia tem se mostrado bastante consistente para discriminar entre modulação positiva e negativa do conjunto de genes representado no microarranjo (Petersen *et al.*, 2005; Twellmeyer *et al.*, 2007; Zaini *et al.*, 2008; Zhang *et al.*, 2008; Fogaça *et al.*, 2010).

Dentre as CDS que não apresentaram correlação entre os perfis de expressão obtidos nos experimentos de microarranjos e ensaios de RT-qPCR (Tabela 10), destacamos que XF0045, XF0794 e XF2637, embora tenham mostrado regulação positiva em resposta ao tratamento com estreptomicina de acordo com dados de microarranjos, os ensaios de RT-qPCR indicam que elas não sejam moduladas (Figura 12). Por outro lado, XF2628 e XFa0052 mostraram perfis de expressão totalmente discrepantes quando seus respectivos dados de microarranjos foram comparados aos de RT-qPCR, onde mostraram modulação negativa no tratamento com estreptomicina (Tabela 10). Esta discrepância pode ser tanto devida a inadequação dos pares de oligonucleotídeos iniciadores confeccionados para o RT-qPCR ou a problemas inerentes a hibridização do fragmento de DNA relativo a estas CDS depositado no microarranjo. Ainda que os pares de oligonucleotídeos para estas 5 CDS mostraram-se específicos em ensaios de PCR de DNA genômico onde amplificaram um único produto no tamanho esperado (dados não mostrados), a modulação destas cinco CDS não será discutida no contexto das alterações no perfil de expressão gênica de *X. fastidiosa* em resposta a exposição a dose subletal de estreptomicina (ver adiante).

Figura 12. Perfis de expressão de CDS selecionadas em resposta ao tratamento com dose subletal de estreptomicina avaliados por RT-qPCR. O RNA total de células tratadas (barras cinzas) ou não (barras pretas) com 1µg/mL de estreptomicina por 15 e/ou 30 min (como indicado ao lado da identificação de cada CDS) foi utilizado para transcrição reversa em cDNA. A quantidade de cDNA de cada amostra foi normalizada pelo gene controle (*dnaQ*; XF2157). A razão de expressão nas amostras não tratadas foi considerada igual a 1. Os dados de RT-qPCR foram obtidos a partir de ensaios realizados em triplicata para duas réplicas biológicas e o gráfico é representativo dos ensaios de uma das réplicas biológicas. A barra de erro indica o desvio padrão médio (SEM).

Tabela 10. Comparação dos dados de microarranjo e de RT-qPCR para 25 CDS.

Identidade do Gene	Produto Predito	Razão de Expressão Microarranjo ^a	Razão de Expressão RT-qPCR ^a
XF0026^b	phospho-2-dehydro-3-deoxyheptonate aldolase	2,03	1,41 ± 0,24
XF0045	transcription-repair coupling factor (ATP-dependent helicase Mfd)	3,01	0,78 ± 0,16
XF0107	30S ribosomal protein S16	2,49	1,79 ± 0,28
XF0121	2-nitropropane dioxygenase (NPD) family protein	0,36/0,40	0,88 ± 0,11
XF0274	6-phosphofructokinase	0,40	0,54 ± 0,14
XF0340	disulfide bond formation protein B	0,42	0,72 ± 0,07
XF0497	hypothetical protein; xfp6	0,31	0,25 ± 0,09
XF0599	catecholate siderophore receptor Fiu (TonB-dependent receptor Fiu)	2,54	3,19 ± 0,61
XF0615	chaperonin GroEL (60kDa chaperonin)	1,55	1,53 ± 0,26
XF0616	co-chaperonin GroES (10 kDa chaperonin)	1,91	1,57 ± 0,4
XF0794	UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate—D-alanyl-D-alanyl ligase	2,16	0,8 ± 0,11
XF1589	plasmid stabilization protein; xfp3	0,26	0,46 ± 0,07
XF1743	esterase; giCVC	3,63	9,58 ± 1,53
XF1827	organic hydroperoxide resistance protein	1,75	1,84 ± 0,32
XF2048	conjugal transfer protein TrbJ	1,62	7,13 ± 1,15
XF2438	30S ribosomal protein S1	2,68	2,11 ± 0,3
XF2466	polygalacturonase (PGA) (Pectinase)	2,22	1,48 ± 0,06
XF2628	elongation factor Tu	1,75	0,1 ± 0,01
XF2637	50S ribosomal protein L11	2,07	0,81 ± 0,15
XF2753	hypothetical protein	0,39/0,33	0,79 ± 0,11/ 0,18 ± 0,07
XFa0052	virulence-associated protein D	2,73	0,66 ± 0,11

^a Os valores da razão de expressão das CDS selecionadas foram extraídos das tabelas 8 e 9 (microarranjos) e de ensaios realizados como descrito na Figura 12 (RT-qPCR).

^b As CDS destacadas em negrito apresentam perfis de expressão coincidentes quando dados de microarranjos e RT-qPCR são comparados.

4.1.1 Análise do perfil de expressão gênica de CDS selecionadas em resposta ao tratamento com dose subletal de estreptomicina

Com o objetivo de refinar os estudos do efeito da exposição de *X. fastidiosa* a dose subletal de estreptomicina, selecionamos um conjunto adicional de 11 CDS para averiguar se sofriam qualquer tipo de modulação. Embora estas CDS não tenham exibido expressão diferencial ou tenham talvez sido excluídas das análises de microarranjos que realizamos, decidimos avaliá-las tendo em vista, que várias CDS deste conjunto, estão potencialmente associadas a processos de patogenicidade e virulência.

Como mostram os resultados apresentados na Tabela 11, a expressão da maioria dessas CDS foi reprimida após o tratamento com 1µg/mL de estreptomicina por 15 min. Destacamos a modulação negativa das CDS XF0287 (*rpfB*) e a ausência de modulação significativa para XF1115 (*rpfF*). Essas CDS codificam fatores de regulação de patogenicidade relacionados à indução da produção do exopolissacarídeo (EPS) que adensa o biofilme de *X. fastidiosa* (Scarpari *et al.*, 2003; Newman *et al.*, 2004; Colnaghi Simionato *et al.*, 2007). Observamos também a modulação negativa da CDS XF0371 (*pilO*), que codifica uma proteína de membrana relacionada à montagem de *pilus* do tipo IV, que está associado a motilidade *twitching* e também à formação de biofilme (De La Fuente *et al.*, 2007; Li *et al.*, 2007). Duas das CDS amostradas XF1171 e XF0760 mostraram modulação positiva nos ensaios de RT-qPCR. As implicações destas observações serão discutidas adiante, conjuntamente com os demais dados de expressão gênica.

Tabela 11. Avaliação do perfil de expressão de 11 CDS utilizando RT-qPCR

Identidade do Gene ^a	Nome do gene ^b	Produto Predito ^b	Razão de Expressão RT-qPCR ^d
XF0169	<i>tyrS</i>	tyrosyl-tRNA synthetase	0,20 ± 0,05
XF0216	-	MarR family transcriptional regulator	0,13 ± 0,04
XF0287	<i>fadD</i> ou <i>rpfB</i>	putative long-chain fatty acyl CoA ligase	0,48 ± 0,04
XF0305	<i>nuoA</i>	NADH dehydrogenase subunit A	0,35 ± 0,08
XF0371	<i>pilO</i>	fimbrial assembly membrane protein	0,13 ± 0,02
XF0760	<i>mutL</i>	DNA mismatch repair protein	3,17 ± 0,24
XF0767	<i>bigR</i>	biofilm growth-associated repressor (BigR) ^c	0,73 ± 0,06
XF0800	<i>ftsQ</i>	cell division protein FtsQ	0,22 ± 0,05
XF0889		Filamentous haemagglutinin (FHA) adhesin like protein	0,64 ± 0,02
XF1115	<i>rpfF</i>	Small diffusible signal molecule regulator (enoyl-CoA hydratase family domain)	1,12 ± 0,21
XF1171	<i>rplO</i>	50S ribosomal protein L15	2,11 ± 0,20

^a Identidade da CDS segundo (Simpson *et al.*, 2000).

^b Nome do gene e o produto predito de acordo com (Simpson *et al.*, 2000) considerando-se, para algumas CDS, a recente re-anotação do genoma da cepa 9a5c disponível em <http://www.xylella.lncc.br>

^c Nome do gene e o produto predito como proposto por (Barbosa e Benedetti, 2007).

^d Os valores da razão de expressão das CDS selecionadas foram determinados por RT-qPCR. O RNA total de células tratadas ou não com 1µg/mL de estreptomicina por 15 min foi utilizado para transcrição reversa em cDNA. A quantidade de cDNA de cada amostra foi normalizada pelo gene controle (*dnaQ*; XF2157). A razão de expressão nas amostras não tratadas foi considerada igual a 1. Os dados de RT-qPCR mostrados foram obtidos a partir de ensaios realizados em triplicata para duas réplicas biológicas. O desvio padrão médio está indicado.

4.1.2. Genes modulados no tratamento de *X. fastidiosa* com dose subletal de estreptomicina

Visando uma análise mais detalhada do conjunto de CDS que sofreram modulação em resposta a dose subletal de estreptomicina, buscamos informações sobre o perfil de expressão destas 136 CDS em outras condições experimentais e descritas na literatura. Esta necessidade nos levou a coletar os dados de expressão publicamente disponíveis para todas as CDS anotadas no genoma de *X. fastidiosa*, os quais foram obtidos com a utilização de microarranjos de DNA. Os dados foram então reunidos visando uma análise integrada da expressão gênica de *X. fastidiosa*, como detalhado no item 4.2. A partir destas informações compusemos as Tabelas 12 e S1 que comparam, respectivamente, o perfil de expressão das CDS moduladas positivamente e negativamente no tratamento com estreptomicina em diversas outras condições experimentais, sendo que estão apresentadas apenas as condições em que foi observada modulação para pelo uma das 136 CDS. A seguir discutiremos as alterações observadas agrupando as CDs segundo suas categorias funcionais ou processos biológicos a que estão relacionadas.

CDS hipotéticas e hipotéticas conservadas

Dentre as 136 CDS consideradas diferencialmente expressas no tratamento de *X. fastidiosa* com 1 µg/mL, 41 estão anotadas como hipotéticas ou hipotéticas conservadas (Tabelas 8 e 9), sendo que apenas duas delas (XF0195 e XF0642) sofreram modulação positiva. XF0195 e XF0642 codificam respectivamente proteínas hipotéticas de 67 e 61 aminoácidos respectivamente, sem ortólogos em outras espécies bacterianas. É interessante observar que XF0195 também sofreu modulação positiva nos estresses salino (exposição a 250mM de NaCl) e osmótico (exposição a 300mM de sacarose) segundo as análises descritas por (Koide, 2006) (Tabela 12).

As CDS XF0766 e XF1243 anotadas como hipotéticas conservadas e negativamente moduladas em resposta ao tratamento com estreptomicina (Tabelas 9 e S1), mostraram níveis de transcritos aumentados tanto no estresse salino como osmótico (Koide, 2006). XF0766 e XF1243 codificam, respectivamente, proteínas hipotéticas de 136 e 399 aminoácidos que são bem conservadas em diversas bactérias.

Merecem destaque, três CDS hipotéticas (XF0890, XF0902 e XF2753) moduladas negativamente no tratamento com estreptomicina, mas com modulação positiva na exposição ao antibiótico gomesina. Estas CDS codificam, respectivamente, proteínas hipotéticas de 283, 175 e 321 aminoácidos, sendo que XF2753 também foi modulada positivamente no estresse osmótico com sacarose (Tabela S1).

Observamos ainda que algumas CDS hipotéticas relacionadas aos fagos de *X.*

fastidiosa (XF0488, XF1661 e XF2514) são reprimidas no tratamento com estreptomicina, uma modulação contrária da verificada nos estresses térmico, salino e osmótico (Tabela S1).

As CDS hipotéticas ou hipotéticas conservadas discutidas acima são, a nosso ver, candidatos interessantes para uma futura caracterização funcional.

CDS relacionadas à resposta de estresse

Alguns estudos demonstraram que exposição a antibióticos que atuam como inibidores da síntese protéica, tais como estreptomicina, puromicina ou canamicina, promove ativação do *regulon* de choque térmico em diversas espécies bacterianas (VanBogelen e Neidhardt, 1990; Mogk *et al.*, 1998; Evers *et al.*, 2001; Ng *et al.*, 2003; Shaw *et al.*, 2003; Qiu *et al.*, 2005). Como já mencionado, a estreptomicina interage diretamente com a subunidade menor do ribossomo, induzindo erros na tradução do mRNA (Mingeot-Leclercq *et al.*, 1999; Carter *et al.*, 2000; Ogle *et al.*, 2003). O aumento na taxa de produção de peptídeos com erros de tradução pode levar à saturação de chaperoninas o que poderia resultar na indução de uma típica resposta de choque térmico (Goff e Goldberg, 1985; Munro e Pelham, 1985; VanBogelen e Neidhardt, 1990; Lund, 2009).

Coincidente com estas observações, nossos resultados de microarranjos e de RT-qPCR (Tabela 10) mostram que a exposição de *X. fastidiosa* a dose subletal de estreptomicina resulta na modulação positiva de XF0615 e XF0616, que codificam respectivamente, a chaperonina de 60 kDa GroEL e a co-chaperonina de 10 kDa GroES (Tabelas 8 e 10). Ambas são induzidas no estresse térmico em *X. fastidiosa* (Koide *et al.*, 2006b) mas não nos estresses osmótico e salino (Koide, 2006). Estas proteínas pertencem à família das proteínas de choque térmico (*heat shock proteins*, *Hsp*) e auxiliam o dobramento adequado de proteínas, sendo que o complexo GroES/GroEL desempenha importante papel em situações que promovem acúmulo de proteínas desenoveladas, como no estresse térmico (Horwich *et al.*, 2007; Lund, 2009).

Além de XF0615 e XF0616, outras 4 CDS (XF0759, XF2048, XFa0025 e XFa0052) positivamente moduladas em resposta a exposição à estreptomicina, também foram induzidas no estresse térmico (Koide *et al.*, 2006b). Destas, destacamos a XF0759 que codifica N-acetilmuramoi-L-alanina amidase ou peptidoglicano amino-hidrolase, uma enzima do grupo das autolisinas e que participa da degradação e remodelamento da parede celular, processos que ocorrem durante a divisão celular ou autólise das células. As autolisinas também podem mediar a liberação de fatores de virulência que não são exportados ativamente (Vollmer *et al.*, 2008). Em conjunto, estas observações embasam futuros experimentos para avaliar se a CDS XF0759 desempenha algum papel na virulência de *X. fastidiosa*.

CDS relacionadas à tradução

O tratamento de *E. coli*, *H. influenzae*, *Streptococcus pneumoniae* e *Y. pestis* com antibióticos que inibem a tradução promove o aumento da síntese de diversas proteínas ribossomais e de fatores de tradução (VanBogelen e Neidhardt, 1990; Evers *et al.*, 2001), além de induzir aumento dos níveis de transcritos desta classe de proteínas (Ng *et al.*, 2003; Qiu *et al.*, 2005). Este mesmo efeito parece ocorrer também em *X. fastidiosa*, uma vez que o tratamento com dose subletal de estreptomicina induziu o acúmulo de transcritos de CDS que codificam as proteínas ribossomais S16, S1 e L11, respectivamente, XF0107, XF2438 e XF2637 (Tabelas 8 e 10). Além disso, ensaios adicionais de RT-qPCR mostraram modulação positiva também para XF1171 que codifica a proteína ribossomal L15 (Tabela 11). Observamos também uma modulação positiva para os fatores de elongação Tu e G (EF-Tu e EF-G), respectivamente codificados pela XF2628 e XF2629, que foi confirmada por RT-qPCR, pelo menos, para esta última CDS (Tabelas 8 e 10). Estas cinco CDS não foram induzidas em nenhuma das condições, além da resposta à estreptomicina, em que o perfil de expressão gênica de *X. fastidiosa* foi avaliado, exceto pela indução de XF2438 em células no biofilme (Tabela 12), o que pode ser um indicativo da regulação específica destas CDS na exposição à estreptomicina.

Para diversas bactérias, está documentado que as proteínas ribossomais regulam sua expressão de modo autógeno, isto é para um dado *operon* de proteínas ribossomais, uma proteína ribossomal específica por ele codificado desempenha papel regulatório, coordenando a expressão de seu próprio *operon* com a disponibilidade dos demais componentes do ribossomo (Zengel e Lindahl, 1994; Allen *et al.*, 2004). Além disso, as proteínas ribossomais regulam sua própria tradução, ligando-se ao seu próprio mRNA e reprimindo sua tradução, como demonstrado para a proteína ribossomal S1 (Boni *et al.*, 2001). Poderíamos especular que o aumento observado nos transcritos para as proteínas ribossomais, em particular a proteína S1, é uma resposta ao bloqueio de sua síntese na presença de estreptomicina, mesmo em doses subletais.

A literatura descreve que a etapa de reconhecimento do aminoacil-tRNA (aa-tRNA) no ribossomo é um processo importante do ciclo de elongação e determina tanto a fidelidade quanto a velocidade da tradução. O aa-tRNA forma um complexo ternário com EF-Tu e GTP ao ser entregue ao sítio A do ribossomo, sendo que o reconhecimento específico códon-anticódon induz uma mudança conformacional no ribossomo, estabilizando a ligação do aa-tRNA e acelerando a reação. Após a hidrólise do GTP, EF-Tu muda sua conformação, se liga a GDP e libera o aa-tRNA. A velocidade de hidrólise do GTP é determinada pela fidelidade da interação códon-anticódon, sendo muito maior na interação do aa-tRNA com

códons cognatos do que não-cognatos (Pape *et al.*, 2000; Ogle *et al.*, 2001; Rodnina e Wintermeyer, 2001; Ogle *et al.*, 2003; Gromadski e Rodnina, 2004a). Na presença de estreptomicina, entretanto, a velocidade de hidrólise do GTP pelo EF-Tu é independente da interação do aa-tRNA ocorrer com códons cognatos ou não-cognatos (Gromadski e Rodnina, 2004b). O aumento nos níveis de EF-Tu (XF2628), detectado na hibridização de microarranjos, possivelmente reflete uma maior necessidade desta proteína para compensar a demanda de EF-Tu ligado a GTP.

A estreptomicina promove uma mudança conformacional no ribossomo que facilita a conversão do estado pós-translocação ao estado pré-translocação, o que pode sinalizar como uma depleção de EF-G (Shoji *et al.*, 2006). Estas observações podem justificar o aumento dos níveis transcricionais de XF2629 (Fator de Elongação G) como uma estratégia para compensar este efeito.

CDS relacionadas à patogenicidade e virulência

Do conjunto de CDS positivamente moduladas no tratamento com estreptomicina, pelo menos uma está diretamente relacionada à patogenicidade em *X. fastidiosa*. Trata-se da CDS XF2466 (*pglA*) que codifica poligalacturonase (Simpson *et al.*, 2000; Van Sluys *et al.*, 2003). A hipótese atual é de que esta enzima, ao lado de outras relacionadas à degradação da parede celular da planta, seja crítica para o sucesso da colonização sistêmica do hospedeiro (Roper *et al.*, 2007b). Entretanto, em cepas isoladas de citros, incluindo a cepa 9a5c, esta CDS possui uma mutação de mudança de fase de leitura (*frameshift*) que não é verificada em cepas isoladas de outras plantas, e que resultaria em uma proteína não-funcional (Van Sluys *et al.*, 2003).

Interessantemente, o tratamento com estreptomicina promoveu a modulação negativa de CDS relacionadas à formação e manutenção de biofilme como a XF0083 (*fimA*), XF2366 (*gumE*) e XF2370 (*gumB*), que codificam respectivamente um precursor da fímbria e duas proteínas responsáveis pela síntese e manutenção do exopolissacarídeo, um importante componente do biofilme de *X. fastidiosa* (Scarpari *et al.*, 2003; de Souza *et al.*, 2004; Newman *et al.*, 2004; Osiro *et al.*, 2004; Souza *et al.*, 2006; Colnaghi Simionato *et al.*, 2007; Li *et al.*, 2007). Por outro lado, o tratamento de *X. fastidiosa* com dose subletal gomesina modula positivamente XF2366 e XF2370, induzindo a formação de biofilme e diminuindo sua virulência (Fogaça *et al.*, 2010). Vale ressaltar que em *Pseudomonas aeruginosa*, concentrações subletais de antibióticos aminoglicosídicos como a estreptomicina induzem a formação de biofilme, o qual contribui para resistência a estes antibióticos (Hoffman *et al.*, 2005).

Coincidente com estas observações, ensaios de RT-qPCR apresentados na Tabela

11 mostraram que o tratamento com 1µg/mL de estreptomicina por 15 min resulta na modulação negativa de XF0287 (*rpfB*) e a ausência de modulação significativa para XF1115 (*rpfF*). XF0287 codifica uma enzima relacionada à síntese de DSF. Esta molécula atua como um sensor de sinalização célula a célula, que é fundamental para a colonização do xilema e controle da virulência da bactéria (Barber *et al.*, 1997; Newman *et al.*, 2004; Chatterjee *et al.*, 2008c). A repressão de XF0287 poderia explicar a repressão das CDS XF2366, XF2370 e XF0083, mencionadas acima e que estão relacionadas com a formação e manutenção do biofilme. Já foi demonstrado que estes genes estão sob controle de uma cascata de sinalização que inclui XF0287 e XF1115, além de outros genes do *operon rpf* (*Regulation Pathogenicity Factors*). Mutantes de *X. fastidiosa* cepa Temecula que não expressam *rpfF* apresentam menor nível de expressão para os genes *gumJ* e *fimA* e são mais virulentos (Chatterjee *et al.*, 2008c).

Observamos ainda a modulação negativa da CDS XF0371 (*pilO*), que codifica uma proteína de membrana relacionada à montagem de *pilus* do tipo IV, que está associado à motilidade *twitching* e também à formação de biofilme (Meng *et al.*, 2005; De La Fuente *et al.*, 2007; Li *et al.*, 2007). Vale ressaltar que esta CDS é positivamente modulada no tratamento com gomesina (Fogaça *et al.*, 2010).

O conjunto destas observações sugere que a exposição à dose subletal de estreptomicina promova um fenótipo de maior virulência, em contraste ao efeito que observamos com o antibiótico peptídico gomesina (Fogaça *et al.*, 2010). O tratamento de *X. fastidiosa* com dose subletal de estreptomicina não promove a formação do biofilme como verificamos recentemente no tratamento com gomesina (Fogaça *et al.*, 2010). A hipótese corrente é que a formação do biofilme é importante nos estágios iniciais da colonização da planta, mas limita sua disseminação sistêmica pela circunscrição da infecção a poucos vasos do xilema (Guilhabert e Kirkpatrick, 2005; Chatterjee *et al.*, 2008a; Chatterjee *et al.*, 2008b; Chatterjee *et al.*, 2008c; Fogaça *et al.*, 2010).

Embora não claramente associada à virulência, vale destacar a modulação positiva de XF0599, que codifica um receptor de sideróforos (Tabela 8). Esta mesma CDS é positivamente modulada em condições de carência de ferro (Tabela 12), possivelmente como uma estratégia para aumentar a captação de ferro complexado a sideróforos. Dados obtidos em nosso laboratório sugerem que a percepção da disponibilidade de ferro parece ser importante na ativação de processos associados à translocação da bactéria através dos vasos do xilema e competição com bactérias endofíticas, mas não para a formação de biofilme (Zaini *et al.*, 2008).

Demais CDS moduladas por estreptomicina

Muitas das CDS moduladas negativamente pelo tratamento de *X. fastidiosa* com estreptomicina estão relacionadas ao metabolismo energético, tais como XF0274, XF305, XF0307, XF1143, XF1855, XF2547 e XF2548 (Tabelas 9 e 11). Uma dramática redução no nível de transcritos de genes relacionados ao metabolismo energético foi observada no tratamento de *Y. pestis* com estreptomicina, porém em concentração muito maior do que a concentração inibitória (Qiu *et al.*, 2005).

A lista de CDS moduladas negativamente por estreptomicina também inclui genes relacionados a processo de reparo de DNA (XF1806, XF2022, XF2552). Por outro lado, observamos a modulação positiva da CDS XF2724, que codifica uma proteína do sistema de restrição do tipo I, e da XF0760 que codifica uma proteína associada a um tipo de reparo de DNA (Tabelas 8 e 11). Observamos também uma expressiva modulação positiva da CDS XF1833, que também ocorre na exposição à gomesina (Tabelas 8 e 12, (Fogaça *et al.*, 2010)). Esta CDS está relacionada à síntese de ubiquinona (Coenzima Q), que além de participar da cadeia de transporte de elétrons, desempenha função antioxidante quando em sua forma reduzida (Do *et al.*, 1996; Poon *et al.*, 2000). Além disso, o tratamento com estreptomicina promoveu a modulação positiva de XF1827, que codifica uma proteína associada à destoxificação de hidroperóxidos orgânicos e, possivelmente, a outros processos como transporte de moléculas orgânicas (Cussiol *et al.*, 2003). Outra CDS que podemos destacar é a XF1743, que codifica uma esterase. Tanto a análise de microarranjos de DNA como o RT-qPCR mostram que os níveis dos transcritos desta CDS são muito aumentados na exposição à estreptomicina. Esta CDS está contida em uma região do genoma descrita como sendo uma ilha genômica (Van Sluys *et al.*, 2003; de Mello Varani *et al.*, 2008).

As listas de CDS moduladas em resposta a exposição de *X. fastidiosa* à dose subletal de estreptomicina (Tabelas 8 e 9) englobam várias outras CDS que não foram aqui discutidas. Procuramos destacar as CDS mais relevantes no contexto da patogenicidade e virulência de *X. fastidiosa*, mas futuros experimentos são necessários para testar as hipóteses derivadas da análise de nossos resultados.

Tabela 12. Comparação do perfil de expressão das CDS moduladas positivamente pela estreptomicina aos seus respectivos perfis quando *X. fastidiosa* é submetida a outras condições experimentais.

Tratam CDS	STR	GOM	HIC	LIC	HS	SS	OS	Glu- 250	BF	PW/ XDM	J1a12	BCYE/ XDM2	PD X1/X2
XF0026	1,0		-3,3	-2,6					1,2				
XF0045	1,6			1,4									
XF0107	1,3					-1,4							
XF0195	1,6					2,4	3,0						
XF0201	1,3						1,4					-1,4	
XF0315	3,4												
XF0461	1,5												
XF0599	1,4		-2,2	1,2									
XF0615	0,6		-5,0		3,9	-2,2	-1,8			-1,1	2,3		
XF0616	0,9		-5,0	-2,1	4,3	-3,2	-2,1			-1,3	2,3		
XF0642	1,2												
XF0647	1,9								0,8				
XF0759	1,5			0,9	1,1					1,2			
XF0794	1,1		1,0										
XF1743	1,9									2,1		-1,2	
XF1827	0,8		-1,8		-0,8								
XF1833	2,2	5,9											
XF2048	0,7		-1,1		0,6								
XF2438	1,4		-1,0	-0,8					0,7				
XF2466	1,2												-0,7
XF2628	0,8			-0,9									
XF2629	1,1		-1,6	-1,4									
XF2637	1,1							-1,0		-1,2			
XF2706	1,3												
XF2724	1,3									1,8			
XFa0025	1,6		1,4		1,1		1,7		1,2				
XFa0052	1,5		1,4		1,4	2,0	3,1		1,1				

As siglas no topo de cada coluna indicam as diferentes condições ou tratamentos em que a expressão gênica de *X. fastidiosa* foi avaliada. Os dados (valor M) da coluna STR foram obtidos como descrito na Tabela 8. Os demais dados (valor M) foram extraídos da literatura como descrito no item 4.2. STR (estreptomicina), GOM (gomesina), HIC (excesso de ferro), LIC (ausência de ferro), HS (choque térmico), SS (sacarose 300mM), OS (NaCl 250mM), Glu-250 (Glicose 250mM), BF (biofilme), PW/XDM (meio PW versus XDM), J1a12 (cepa J1a12 versus 9a5c), BCYE/XDM2 (meio BCYE versus XDM2) e PD X1/X2 (cepa PD xilema de videira versus de citros).

4.2. Análise integrada de perfis de expressão gênica global de *Xylella fastidiosa*

Após o sequenciamento completo do genoma de *X. fastidiosa* cepa 9a5c (Simpson *et al.*, 2000), diversos outros estudos se seguiram, incluindo a elucidação completa ou parcial de genomas de outras cepas de *X. fastidiosa* e avaliações do transcrito de *X. fastidiosa* em diversas condições experimentais. Estas informações estão disponíveis em bancos de dados públicos, em publicações científicas, dissertações e teses, mas não estão organizadas de modo a possibilitar identificação de correlações entre estes dados, em particular aqueles relativos à expressão gênica em *X. fastidiosa* obtidos através da análise de microarranjos de DNA. Evidentemente, que este tipo de análise integrada é uma prospecção de correlações e requer validação experimental posterior, uma vez que são desconsiderados detalhes metodológicos das condições experimentais amostradas e diferenças tanto entre plataformas de microarranjos de DNA como nas estratégias de análise de dados.

Para atender nosso interesse em comparar o perfil de expressão gênica de *X. fastidiosa* em resposta ao tratamento com estreptomicina com perfis de expressão observados em outras condições experimentais (item 4.1.2.), recuperamos manualmente dados de expressão gênica diferencial para as CDS anotadas no genoma de *X. fastidiosa* cepa 9a5c e para suas ortólogas na cepa Temecula, os quais estão publicamente disponíveis. As Tabelas 13 e S2 (Anexos) mostram, respectivamente, a descrição resumida das condições experimentais e os dados que foram manualmente extraídos das fontes citadas, sendo que dados de expressão apresentados como razão de expressão foram convertidos para $\log_2$ (razão de expressão). Embora os valores de expressão diferencial anotados na planilha (Tabela S2, Anexos) correspondam fielmente à fonte original de cada conjunto de dados, os valores absolutos não são comparáveis entre si, pois tanto as plataformas de microarranjos como as estratégias de análise diferem em cada caso. Além disso, a planilha em seu formato atual, não informa se uma CDS foi expressa naquela condição, isto é se foi detectada na análise de microarranjos, ainda que não tenha sido diferenciantemente expressa. Planejamos, futuramente, recuperar tal informação do banco de dados GEO (<http://www.ncbi.nlm.nih.gov/geo/>), quando disponível, ou diretamente junto aos autores dos artigos. Infelizmente, nem todas as fontes que consultamos seguiram as normas do MIAME (*Minimum Information About a Microarray Experiment*) que são recomendações sobre como reportar experimentos com microarranjos e não depositaram os dados no GEO (Brazma *et al.*, 2001; Edgar e Barrett, 2006).

Mesmo com estas ressalvas, os dados já compilados possibilitam a obtenção de agrupamentos hierárquicos, como o que está mostrado na Figura 13. Para confecção deste

heatmap a escala de valores de $\log_2$ da razão de expressão foi fixada arbitrariamente de -7 a +7 como limites mínimo e máximo apenas para CDS com expressão diferencial. No tamanho apresentado, a figura não tem resolução para distinguirmos cada uma das 1472 CDS moduladas em pelo menos uma das 9 condições experimentais incluídas nesta análise, entretanto, o *heatmap* mostra claramente que estes os perfis de expressão exibem mais diferenças do que semelhanças, ainda que 65 de 1472 CDS tenham sido moduladas positivamente em pelo menos 3 das 9 condições amostradas.

Considerando as 20 condições experimentais incluídas na análise integrada e listadas na Tabela 13, observamos que 1859 das 2848 CDS anotadas no genoma de *X. fastidiosa* cepa 9a5c foram moduladas em pelo menos uma condição. Além daquelas com expressão modulada por estreptomicina e discutidas anteriormente (Tabelas 10 e 12), apontamos algumas correlações que poderão ser exploradas futuramente visando uma caracterização mais detalhada de regulação de expressão e/ou para identificação de alvos para análise funcional, sendo que excluímos CDS relacionadas aos fagos contidos no genoma da cepa 9a5c. A Tabela S3 (Anexos) lista CDS selecionadas, por exemplo: a) CDS do operon de síntese de EPS (XF2364, XF2367 e XF2369) que, dentre as condições listadas, são moduladas positivamente apenas em células expostas a dose subletal de gomesina e no mutante do regulador transcricional *gacA*. Recentemente foi mostrado que CDS deste operon estão entre o conjunto de genes modulados¹ positivamente pela adição de pectina e glucana ao meio de cultivo (Killiny e Almeida, 2009a); b) CDS associadas à motilidade do tipo *twitching* (XF1632 e XF 1633) são moduladas negativamente em condições em que a concentração de ferro é alterada, porém positivamente quando a cepa Temecula é cultivada na presença de xilema isolado de videira e c) CDS XF2196 (proteína similar a hemaglutinina) é modulada positivamente quando *X. fastidiosa* é exposta à gomesina, ao estresse térmico ou a condição limitante de ferro. Além destas, algumas CDS com função desconhecida (XF0449, XF0989, XF1755, XF1756, XF2177, XF2195 e XF2257) e com expressão modulada em mais de uma condição experimental são, a nosso ver, bons candidatos para uma caracterização bioquímica e funcional. Interessantemente, a ortóloga da CDS XF2195 está entre as CDS da cepa Temecula que são moduladas positivamente por glucana, um carboidrato do hospedeiro vegetal que estimula a transmissão da bactéria para o inseto (Killiny e Almeida, 2009a).

¹ A lista de genes de *X. fastidiosa* modulados por carboidratos do hospedeiro vegetal descrita por Killiny e Almeida (2009) ainda não foi incorporada em nossa análise integrada.

Tabela 13. Condições experimentais incluídas na análise integrada de expressão gênica de *X. fastidiosa*

Sigla	Condição Experimental	Plataforma – Microarranjo ¹	Cepa de <i>X. fastidiosa</i>	Referência
HS	Choque térmico 40°C (7 a 45 min) em meio PW	CAGE/IQUSP	9a5c	(Koide <i>et al.</i> , 2006b)
OS	Exposição a NaCl 250mM (7 a 60 min) em meio PW	CAGE/IQUSP	9a5c	(Koide, 2006)
SS	Exposição a sacarose 300mM (7 a 60 min) em meio PW	CAGE/IQUSP	9a5c	(Koide, 2006)
J1a12	Comparação J1a12 e 9a5c cultivo em meio PW	CAGE/IQUSP	J1a12 e 9a5c	(Koide <i>et al.</i> , 2004b)
Glu1	Cultivo em meio XDM com Glicose 1mM x 50mM	CAGE/IQUSP	9a5c	(Pashalidis <i>et al.</i> , 2005)
Glu2	Cultivo em meio XDM com Glicose 250mM x 50mM	CAGE/IQUSP	9a5c	(Pashalidis <i>et al.</i> , 2005)
HIC	Exposição a excesso de ferro (15 a 60 min)	CAGE/IQUSP	9a5c	(Zaini <i>et al.</i> , 2008)
LIC	Exposição a carência de ferro (15 a 60 min)	CAGE/IQUSP	9a5c	(Zaini <i>et al.</i> , 2008)
GOM	Tratamento com dose subletal de gomesina (60 min)	CAGE/IQUSP	9a5c	(Fogaça <i>et al.</i> , 2010)
STR	Tratamento com dose subletal de estreptomicina (15 a 60 min)	CAGE/IQUSP	9a5c	Este trabalho
Rpo_N	Mutante $\Delta rpoN$ ($\sigma 54$)	CAGE/IQUSP	J1a12 e 9a5c	(da Silva Neto <i>et al.</i> , 2008)
Rpo_E	$\Delta rpoE$ submetido a choque térmico 40°C (25 min) (σE)	CAGE/IQUSP	J1a12	(da Silva Neto <i>et al.</i> , 2007)
FP	Células de primeira passagem x células de 46 passagens em meio PW	UMC	9a5c	(de Souza <i>et al.</i> , 2003)
BF	Células de biofilme x células planctônicas <i>in vitro</i>	UMC	9a5c	(de Souza <i>et al.</i> , 2004)
BF2	Células de biofilme x células planctônicas <i>in vitro</i> na presença de lâminas de lenho	UMC	9a5c	(Caldana, 2002)
XDM/PW	Cultivo em meio XDM e PW	UMC	9a5c	(Nunes <i>et al.</i> , 2003)
BCYE/XDM2	Cultivo em meio BCYE e XDM2	UNESP	9a5c	(Travensolo <i>et al.</i> , 2009)
gacA_PD	Mutante $\Delta gacA$	NimbleGen	PD A05	(Shi <i>et al.</i> , 2009b)
algU_PD	Mutante $\Delta algU$	NimbleGen	PD A05	(Shi <i>et al.</i> , 2007)
PD X1/X2	Cultivo em xilema de videira versus citros	Macroarranjo em náilon	PD A05	(Shi <i>et al.</i> , 2009a)

¹CAGE/IQUSP: Projeto Cooperação para Análise da Expressão Gênica no IQUSP

UMC: Universidade de Mogi das Cruzes

UNESP: Universidade do Estado de São Paulo- campus Jaboticabal

NimbleGen: <http://www.nimblegen.com/>

Figura 13. Comparação de perfis de expressão *Xylella fastidiosa* submetida a diferentes condições experimentais. O *heatmap* foi gerado utilizando-se um programa (*script*) doméstico escrito em R e os valores M da Tabela S2. Como há pouca intersecção entre os experimentos (sempre há algum gene que está em um experimento, mas não no(s) outro(s)), a medida de similaridade que geralmente é usada não é bem definida e o algoritmo de *clustering* hierárquico falha. Desta forma os genes foram arbitrariamente ordenados para geração do *heatmap*.

4.3. Sequenciamento do genoma da cepa J1a12 de *Xylella fastidiosa*

Está em andamento em nosso grupo de pesquisa, um projeto cujo propósito é a identificação de novos genes potencialmente associados a aspectos evolutivos, à adaptação e à virulência de *X. fastidiosa* através do sequenciamento e análise comparativa de genomas de outras cepas deste fitopatógeno. Foram selecionadas cepas isoladas de citros, de cafeeiros, de ameixeiras e de uma planta ornamental (*Hibiscus*) para o sequenciamento genômico utilizando-se a plataforma de pirosequenciamento 454 (454 Life Sciences/Roche). Dentre as cepas isoladas de citros selecionadas para este estudo está a J1a12, uma cepa não-virulenta. Como parte deste trabalho, nós realizamos sequenciamento do genoma desta cepa e a análise inicial das sequências já obtidas.

4.3.1. Confirmação do fenótipo não-virulento da cepa J1a12.

A cepa J1a12 foi isolada de laranjeiras com sintomas de CVC. Até o presente, das cepas isoladas de citros, apenas a J1a12 é comprovadamente suscetível a manipulação genética (Monteiro *et al.*, 2001b; da Silva Neto *et al.*, 2002; Koide *et al.*, 2004a). Entretanto, foi demonstrado que esta cepa exibe fenótipo menos virulento em *Nicotiana tabacum* e *Citrus sinensis* em comparação à cepa 9a5c (Koide *et al.*, 2004b). Vale ressaltar que a virulência da cepa 9a5c já foi confirmada em diversos hospedeiros (Li *et al.*, 1999; Lopes *et al.*, 2000; Monteiro *et al.*, 2001a; Alves *et al.*, 2003; de Souza *et al.*, 2003; Koide *et al.*, 2004b).

Estas observações foram reproduzidas em dois experimentos independentes realizados por Andréa C. Fogaça em colaboração com Diva C. Teixeira e Nelson Wulff (Fundecitrus, Araraquara, SP), onde não foram observadas lesões foliares nas plântulas de tabaco (*Nicotiana glauca*) inoculadas com a cepa J1a12, sendo que estas plantas mantiveram-se saudáveis por 7 meses enquanto que aquelas infectadas com 9a5c sucumbiram completamente à infecção (dados não mostrados).

Nós repetimos os ensaios de infecção em colaboração com Diva C. Teixeira, sendo que incluímos nestes experimentos a co-infecção das plantas com as cepas J1a12 e 9a5c. O aparecimento de lesões foliares típicas de CVC foi avaliado após 5, 7, 9 e 12 semanas, sendo que após 12 semanas foram coletadas folhas para extração de DNA e realização de ensaios de PCR para confirmar ou não a presença de DNA genômico de *X. fastidiosa*, que é indicativo da colonização da planta. Os resultados sumarizados na Tabela 14 confirmam que a cepa J1a12 não é virulenta quando comparada a cepa 9a5c. Os resultados também mostram que o fenótipo virulento exibido pela cepa 9a5c sofreu modesta redução quando o inóculo foi realizado com células de 26 passagens em cultivo. Já foi demonstrado que a

cepa 9a5c com número maior de passagens (46 passagens em cultura por um ano) exibe redução significativa da virulência (de Souza *et al.*, 2003).

O inóculo com uma mistura equivalente de J1a12 e 9a5c, como esperado resultou em plantas sintomáticas (Tabela 14), entretanto não foi possível neste experimento avaliar as proporções das duas cepas após 12 semanas. Em um experimento similar realizado anteriormente no nosso laboratório, foi observado que a colonização de plantas pela cepa J1a12 de *X. fastidiosa* parece ser favorecida pela presença de 9a5c, tendo sido observado um aumento de 14% de detecção de J1a12 em amostras de DNA de plantas co-inoculadas em comparação com amostras de DNA de plantas inoculadas apenas com J1a12 (dados não apresentados).

A comparação entre genomas das cepas J1a12 e 9a5c utilizando microarranjos de DNA revelou diferenças significativas que podem estar correlacionada com o fenótipo menos virulento da cepa J1a12 (Koide *et al.*, 2004b). Foi verificada ausência ou alta divergência de sequência em 14 genes na cepa J1a12, além de divergência de sequência de outros 50 genes. Outro subconjunto de genes está aparentemente em maior número de cópias na cepa J1a12 em relação à cepa 9a5c. O subconjunto de genes ausentes ou altamente divergentes na cepa J1a12 inclui as CDS da arginase (XF1250) e da adesina da fimbria (XF0077). O subconjunto de 50 genes divergentes inclui as CDS XF0078, XF1968, XF2542 e XF2726, que codificam respectivamente adesina da fimbria, metiltransferase, proteína da fimbria e uma proteína do sistema de restrição-modificação do tipo 1.

Visando aprofundar a comparação dos genomas das cepas J1a12 e 9a5c e identificar outras diferenças não detectadas por microarranjos de DNA, realizamos o sequenciamento do genoma da cepa J1a12.

Tabela 14. Avaliação da virulência de *Xylella fastidiosa* cepas 9a5c e J1a12 inoculadas em *Nicotiana clevelandii*

		Plantas sintomáticas ^c /Total de plantas vivas no tempo indicado após a inoculação				Plantas positivas no PCR ^d /Total de plantas vivas após a inoculação
Inóculo ^a	Passagens em cultivo ^b	5	7	9	12	12
		Semanas				semanas
9a5c (30)	17	3/29 (10,3%)	12/28 (42,9%)	13/26 (50%)	20/25 (80%)	22/25 (88%)
J1a12 (30)	13	0/29 (0%)	0/28 (0%)	0/28 (0%)	0/27 (0%)	14/27 (51,8%)
9a5c (12)	26	1/12 (8,3%)	4/12 (33,3%)	5/12 (41,7%)	8/11 (72,7%)	9/11 (81,8%)
J1a12 (12)	26	2/11 (18,2%)	0/10 (0%)	0/10 (0%)	0/10 (0%)	5/10 (50%)
9a5c + J1a12 (30)	17 13	0/30 (0%)	9/29 (31%)	11/29 (37,9%)	22/29 (75,9%)	27/29 (93%)
PBS ^e (12)	-	0/11 (0%)	0/10 (0%)	0/10 (0%)	0/9 (0%)	0/9 (0%)
Controle ^f (16)	-	0/16 (0%)	0/16 (0%)	0/16 (0%)	0/12 (0%)	0/12 (0%)

^aO número de plantas sadias inoculadas está indicado entre parênteses.

^bNúmero de passagens das respectivas cepas em cultivo no laboratório em meio completo PW.

^cOs sintomas foram avaliados pela presença de lesões foliares nas folhas da planta.

^dOs ensaios de PCR foram realizados em amostras das plantas utilizando-se iniciadores específicos para isolados de citros (Pooler e Hartung, 1995).

^eInóculo de solução salina tamponada.

^fPlantas sadias não submetidas a nenhum inóculo.

4.3.2. Pirosequenciamento e montagem dos genomas das cepas 9a5c e J1a12.

O pirosequenciamento do genoma de *X. fastidiosa* cepa J1a12 foi realizado a partir de 5µg de DNA genômico com alto grau de pureza e de boa qualidade, obtidos de dois cultivos independentes de J1a12, sendo um deles com 15 passagens (15P) e o outro com 24 passagens (24P). Tendo em vista que a metodologia está em fase de implantação no Centro Avançado de Tecnologias em Genômica (CATG) do IQUSP, optamos por re-sequenciar o genoma da cepa 9a5c para efeito de comparação. O DNA desta cepa foi preparado a partir de células com 14 passagens em cultivo (14P). Os sequenciamentos foram realizados como descrito no item 3.9., utilizando-se reagentes da série titânio (GS-FLX Titanium Series Reagents) e o equipamento 454 Genome Sequencer FLX (454 Life Sciences/Roche), instalado no CATG. A tabela 15 resume as informações sobre o sequenciamento destas 3 amostras. A diferença de rendimento entre as três preparações possivelmente é reflexo de diferenças na qualidade do DNA genômico de partida. Ainda assim, nas 3 amostras o rendimento obtido foi abaixo do anunciado pelo fabricante, que é de ~250.000 *reads* com tamanho médio de 400 bases por região da placa quando esta é dividida em 4 regiões. A figura 14 mostra a distribuição de *reads* por tamanho referente ao sequenciamento das amostras de DNA das cepas J1a12 15P, J1a12 24P e 9a5c. Os perfis obtidos são ligeiramente distintos do anunciado pelo fabricante, que reporta um número menor de *reads* com tamanho abaixo de 400 bases e uma distribuição mais equilibrada próxima ao tamanho de 500 bases. Esta diferença pode ser devida a problemas técnicos associados ao sequenciador e não necessariamente à qualidade das amostras sequenciadas (S. Verjovski-Almeida, comunicação pessoal).

As sequências obtidas foram submetidas ao programa *Newbler* (454 Life Sciences/Roche) para uma montagem preliminar, cujo resultado está apresentado na Tabela 16. Foram feitas variações nas estratégias de montagens utilizando-se a plataforma SABIA (*System for Automated Bacterial (genome) Integrated Annotation*) (Almeida *et al.*, 2004) que integra uma série de programas de montagem entre eles o próprio *Newbler*. Além disso, foi realizado o mapeamento das sequências geradas no genoma referência da 9a5c (Simpson *et al.*, 2000) com esta mesma plataforma.

Os resultados mostraram que os dois plasmídeos da cepa 9a5c foram sequenciados, sendo que falta 1 base para completar a sequência do plasmídeo de 1.285 pb (pXF1.3) enquanto que o plasmídeo pXF51 de 51.158 tem uma lacuna de 100pb. O mapeamento dos *contigs* no genoma referência revela que as sequências faltantes estão dispersas e correspondem a regiões pequenas, sendo que a maior região que não foi sequenciada ou montada corretamente equivale a 15.000 pb. A identidade observada entre as sequências de 9a5c que geramos e o genoma referência é de 100%, como esperado.

As sequências das cepas J1a12 15P e 24P foram montadas separadamente e também em conjunto, utilizando apenas o programa Newbler (Tabela 16) ou a plataforma SABIA (*System for Automated Bacterial (genome) Integrated Annotation*) (Almeida *et al.*, 2004). O resultado das montagens individuais não mostrou diferenças entre as duas amostras de J1a12, sendo que a montagem combinando os *reads* de ambas as amostras resultou em 374 *contigs* compreendendo o total de 2.614.884 pb, um resultado melhor do que as montagens independentes. Dos 374 *contigs* de J1a12, 238 *contigs* (2.483.197 pb) mapearam no cromossomo do genoma referência de 9a5c. Interessantemente, 43 *contigs*, os quais totalizam 57.521 pb, não mapearam no genoma de 9a5c e apontam para regiões exclusivas de J1a12 em relação a 9a5c, como será discutido adiante. Por outro lado, 195.549 pb da cepa 9a5c estão ausentes no genoma de J1a12. A distribuição das sequências de J1a12 no genoma referência está resumida na Tabela S5.

Não há evidência para existência do plasmídeo menor (pXF1.3) na cepa J1a12 (Koide *et al.*, 2004b), e concordante com esta observação nenhuma sequência ou *contig* de J1a12 mapeou na sequência de pXF1.3. Por outro lado, um *contig* de 51.126 pb de J1a12 mapeou na sequência de pXF51 da cepa 9a5c, o qual está descrito com tamanho de 51.158 (Simpson *et al.*, 2000).

Em resumo, as análises já realizadas indicam que o sequenciamento realizado não foi suficiente para obtermos o genoma completo de J1a12, uma vez que os *reads/contigs* obtidos cobrem ~93% da sequência do genoma de 9a5c. Outras estratégias terão que ser utilizadas para o fechamento das lacunas existentes e já identificadas (Tabela S5). Uma observação importante é que as regiões repetidas do genoma, principalmente as regiões dos pró-fagos, interferem na montagem e irão requerer experimentos adicionais para serem resolvidas (Van Sluys *et al.*, 2003; de Mello Varani *et al.*, 2008).

Tabela 15. Desempenho do pirosequenciamento de DNA cepas J1a12 e 9a5c

Cepa ^a	Número de leituras (total)	Tamanho médio das leituras (bases)	Total de bases sequenciadas	Cobertura estimada do genoma ^b
J1a12 (15)	206.104	401	~ 71 x 10 ⁶	~26 vezes
J1a12 (24)	94.765	360	~ 31 x 10 ⁶	~11 vezes
9a5c (14)	132.701	386	~ 45 x 10 ⁶	~16 vezes

^aO número de passagens do cultivo em cultura está indicado entre parênteses.

^b Considerando-se o tamanho do genoma da cepa 9a5c (~2,7 x 10⁶ pb)

Figura 14. Distribuição das seqüências geradas no sequenciador 454 por tamanho em bases. A figura é uma composição de imagens recuperadas da tela do computador que gerencia o sequenciador 454 e referem-se às 3 regiões da placa em que foram aplicadas as bibliotecas de J1a12 15P, J1a12 24P e 9a5c 15P, respectivamente, região 1, 2 e 4. A região 3 continha uma amostra não relacionada. Os gráficos mostram o tamanho das seqüências (em bases) no eixo x e número de seqüências no eixo y. As informações sobre o tamanho médio dos *reads*, maior e menor *read* por região e total de *reads* que passaram os critérios de qualidade estão mostradas na caixa cinza.

Tabela 16. Montagem das sequências das cepas J1a12 e 9a5c^b

Cepa ^a	Número de sequências utilizadas na montagem	Número total de <i>Contigs</i>	Número de <i>Contigs</i> maiores	Tamanho do maior <i>Contig</i> (bases)	Total de bases na montagem
J1a12 (15)	204.095	548	395	78.483	2.582.157
J1a12 (24)	93.386	1.319	1.005	24.548	2.265.861
J1a12 15P + 24P	297.481	-	374	115.071	2.614.884
9a5c (14)	130.931	465	371	87.084	2.411.891

^aO número de passagens está indicado entre parênteses.

^b *Contigs* montados pelo programa *Newbler* ou plataforma SABIA

4.3.3. Comparação preliminar do genoma de J1a12 com o genoma de 9a5c.

Demos início à comparação do genoma parcialmente sequenciado de J1a12 com o genoma completo de 9a5c, através de análises manuais utilizando BLAST (*Basic Local Alignment Search Tool*) e ferramentas de comparação e anotação de sequências de DNA e de proteínas disponíveis no servidor RAST (Aziz *et al.*, 2008). Inicialmente pesquisamos as diferenças entre estes dois genomas previamente apontadas pela análise de microarranjos de DNA (Koide *et al.*, 2004b). Em todos os casos as divergências anteriormente identificadas foram confirmadas, como ilustrado por alguns exemplos (Tabela 17).

Destacamos a aparente ausência no genoma de J1a12 de uma das cópias das CDS que codificam precursores de microcinas similares às colicinas V. As duas cópias mantidas em J1a12 são similares, mas não idênticas as suas ortólogas na cepa 9a5c (Figura 15). As demais CDS relacionadas à síntese e secreção das colicinas V bem como de imunidade à toxina (Simpson *et al.*, 2000; Pashalidis *et al.*, 2005) são idênticas nos dois genomas. As colicinas V são peptídeos antimicrobianos secretados por alguns membros de Enterobacteriaceae que atuam contra bactérias sensíveis e são importantes na adaptação da bactéria ao seu *habitat* (Havarstein *et al.*, 1994). Em *X. fastidiosa*, pelo menos uma CDS codificadora da colicina V tem níveis de transcritos maiores em bactérias recém-isoladas de plantas, seguindo-se dramática diminuição em bactérias mantidas em cultura axênica (de Souza *et al.*, 2005). Ainda que não exista evidência bioquímica de que destes polipeptídeos sejam efetivamente produzidos durante a colonização do hospedeiro pela *Xylella fastidiosa*, é interessante especular se a ausência de uma das CDS na cepa J1a12 contribuiria para

sua menor capacidade de colonização do hospedeiro vegetal.

A sequência do genoma de J1a12, mesmo que incompleta, revelou a presença de todos os genes encontrados nos *operons gum* e *rpf* de *X. fastidiosa*, os quais apresentam alta similaridade com os respectivos genes da cepa 9a5c (Simpson *et al.*, 2000; da Silva *et al.*, 2001). Estas observações são indicativas de que a cepa J1a12 possui toda a maquinaria necessária para a biossíntese do exopolissacarídeo (EPS) bem como para regulação da sua expressão. Como previamente observado (Koide *et al.*, 2004b), a incapacidade da cepa J1a12 de produzir e manter um biofilme denso deve-se provalmente à ausência de uma (XF0077) das 3 CDS que codificam adesina da fímbria. Os dados de sequenciamento confirmaram a deleção desta CDS e mostraram que as outras duas, XF0078 e XF0080, estão presentes, respectivamente com 86% e 100% de identidade em relação às sequências de aminoácidos de suas ortólogas na cepa 9a5c. Nossas análises iniciais não identificaram diferenças importantes no genoma de J1a12 com relação aos genes associados à síntese do *pilus* do tipo IV, o que nos leva a supor que embora a capacidade de adesão de J1a12 seja diminuída, esta cepa provavelmente deve ser capaz de realizar a motilidade do tipo *twitching*. Já foi demonstrado que, em *X. fastidiosa* cepa Temecula, o *pilus* do tipo IV está associado a motilidade *twitching* além de também contribuir para formação do biofilme (Meng *et al.*, 2005; De La Fuente *et al.*, 2008).

Por outro lado, as sequências obtidas até o momento, sugerem o genoma de J1a12 possui pelo menos duas CDS similares às adesinas do tipo das hemaglutininas codificadas no genoma de 9a5c pelas CDS XF0889, XF2196 e XF2775. Estas regiões do genoma oferecem problemas de montagem em virtude de suas sequências repetitivas, o que tornam prematuras conclusões definitivas quanto à ausência efetiva de uma destas CDS em J1a12.

Uma observação interessante refere-se às sequências que aparentemente são exclusivas de J1a12, isto é *contigs* que não mapearam no genoma de 9a5c. Alguns *contigs* apresentam segmentos com baixa identidade ao genoma de 9a5c ou de outras cepas de *X. fastidiosa*, sendo que algumas destas são sugestivas de regiões possivelmente duplicadas no genoma de J1a12. Entretanto, um dos 43 *contigs* (*contig* 374 de 27.258 pb) não alinhou com o genoma de 9a5c. Buscas no *GenBank* [base dados nr/nt (nucleotide collection) e wgs (*whole genome shot gun reads*)] utilizando a ferramenta BLASTn revelaram que o *contig* 374 apresenta regiões com identidade a sequências genômicas de *X. fastidiosa* cepa Ann-1 (dados não apresentados) e com o plasmídeo pVEIS01 de *Verminephrobacter eiseniae* que possui 31.194 pb (Figura 16). A maior região de similaridade entre o *contig* 374 e sequências de *X. fastidiosa* Ann-1 apresenta ~13300 pb, com 94% de identidade e está compreendida entre as posições de ~11670 e ~24973 (dados não apresentados). A figura 16 mostra o resultado do BLASTn esquematizando as regiões de similaridade com pVEIS01, as quais correspondem a 66% do *contig* 374 e apresentam até 84% de identidade. O alinhamento

das sequências do *contig* 374 e de pVEIS01 está apresentado na figura S1, sendo que a região que apresenta maior identidade está compreendida entre a posição ~2.700 e ~15.200 de pVEIS01. *Verminephrobacter eiseniae* é uma bactéria Gram negativa, simbiote de minhocas (*Eisenia foetida*) (Pinel *et al.*, 2008), cuja sequência genômica está publicamente disponível. Vale destacar que há um relato descrevendo a existência de um plasmídeo de *X. fastidiosa* isolada de amoreira, o qual apresenta extensa similaridade ao plasmídeo de *V. eiseniae* (Stenger e Chen, 2008), porém sua sequência não está publicamente disponível. Observamos também que o *contig* 374 apresenta um segmento de 761 pb com 97% de identidade a uma região compreendida entre a posição 22.031 a 22791 do pXF51, um plasmídeo originalmente descrito na cepa 9a5c (Simpson *et al.*, 2000) e que também está presente na cepa J1a12.

Análises utilizando as ferramentas BLASTx e PSI-BLAST associadas a anotação automática realizada no servidor RAST sugerem a organização apresentada na figura 17 para as CDS preditas a partir da sequência do *contig* 374 que inclui genes *VirB/VirD*, os quais são relacionados ao sistema de secreção do tipo 4 em bactérias (Alvarez-Martinez e Christie, 2009). Interessantemente, algumas das CDS do *contig* 374 apresentam maior similaridade de sequência de aminoácidos com CDS ortólogas existentes no cromossomo ou no plasmídeo de *X. fastidiosa* cepa Ann-1 do que de outras cepas (Bhattacharyya *et al.*, 2002a), sendo que a organização dos genes relembra a encontrada no plasmídeo pXF51 existente na cepa 9a5c (Marques *et al.*, 2001). Além das CDS que codificam a toxina PemK e antitoxina PemI relacionadas a função de manutenção de plasmídeo e também codificadas no pXF51 (Marques *et al.*, 2001), a anotação preliminar do *contig* 374 identificou uma CDS que codifica uma provável toxina similar à toxina zeta (CDS 18). Esta toxina atua como inibidora do crescimento bacteriano, quando na ausência da antitoxina épsilon (Lioy *et al.*, 2010). Especulamos que a CDS hipotética vizinha a CDS 18 possa codificar uma ortóloga da antitoxina épsilon. Análises e experimentos adicionais são necessários para obter a sequência completa deste novo plasmídeo revelado pelo *contig* 374 e se a provável toxina zeta é funcional na cepa J1a12. Será interessante também avaliar a presença deste novo plasmídeo em outros isolados de citros.

As análises comparativas do genoma de J1a12 serão futuramente refinadas. Contudo, os resultados que obtivemos até o momento confirmam os dados anteriores sobre diferenças encontradas neste genoma relativamente ao genoma da cepa 9a5c (Koide *et al.*, 2004b) e apontam para diferenças adicionais que podem contribuir para explicar as características fenotípicas da cepa J1a12.

Tabela 17. Comparação entre genomas de J1a12 e 9a5c: CDS previamente identificadas como altamente divergentes ou divergentes

		Resultado da comparação usando
CDS	Produto predito	BLAST
XF0075	Hypothetical protein	CDS incompleta
XF0077	Fimbrial adhesion precursor	CDS com deleção
XF0078	Fimbrial adhesion precursor	86% identidade (sequência de aminoácidos)
XF0496 a XF0502	Hypothetical proteins; xfp6	ausente
XF0262	Colicin V precursor	Apenas 2 cópias destas CDS em J1a12
XF0263	Colicin V precursor	
XF0264	Colicin V precursor	
XF0663 a XF0667	Hypothetical proteins	J1a12 apresenta deleções nesta região, sendo que XF0668 aparece truncada
XF0668	hemolysin-type calcium binding protein	
XF0684	phage-related protein ; xfp1	CDS com polimorfismos
XF0696	phage-related repressor protein ; xfp1	CDS com polimorfismos
XF1250	Arginine deaminase	ausente
XF1646	UDP-3-O-(R-3-hydroxymyristoyl)-glucosamine N-acyltransferase	ausente
XF1609	glucose/galactose transporter	94% identidade (sequência de aminoácidos)
XF1707	Hypothetical protein ; xfp4	ausente
XF1708	Putative HTH-type addiction module antidote protein, HigA family; xfp4	ausente
XF1874	Hypothetical protein	CDS com deleção
XF1884	Hypothetical protein	CDS com deleção
XF1968	Type III restriction-modification system Mod protein (adenine specific methyltransferase)	CDS com polimorfismos
XF2407	Bacteriocin	J1a12 apresenta deleções e poliformismos nesta região genômica
XF2542	Tfp pilus assembly protein, major pilin FimA/PilA (Type IV fimbrellin)	67% identidade (sequência de aminoácidos)
XF2722	Type I restriction-modification enzyme, S subunit (DNA specificity domain)	69% identidade (sequência de aminoácidos)
XF2726	Type I restriction-modification enzyme, S subunit (DNA specificity domain)	36% identidade (sequência de aminoácidos)
XFb0001	Replication protein	ausente
XFb0002	Hypothetical protein	ausente

```

XF0263      MRELTSIEMNNVSGGD-LATRIEASIVFGVSAFFAGSIWGGTRGGDGGGILGVGSIAQGV 59
1_J1a12     MRELTSIEMNNVSGGD-LASRLEASIVFGVSAFFAGSIWGGARGGDGGGILGIGSIGQGV 59
XF0264      MRELTSIEMNNVSGGD-FATRLEASIVGGIAAFFAGSIWGGTRGGDGGGVLGIGSIGQGV 59
XF0262      MRELTLTEIDNVSGAD-LGSRLSAAIVGGVAAFFAGSIWGGTRGGDGGGILGVGSIGQGV 59
2_J1a12     MRELTLTEINNVSGGTTLASRLEGALVWGINAGFVGSVWGGTRGGDGGGILGIGSIGQGV 60
          *****  *::*****.  .:::..::*  *:  *  *.***:***:*****:***:***.***

XF0263      GMVYGGIVGGIGGLIAGFVLDKNVTYNYAVGFYNSLFNGTFTK 102
1_J1a12     GMVYGGIVGGIGGLVAGFVLDKTVTYNYAVGFYNSLFNGTFAK 102
XF0264      GMVYGGIAGAIGGAIGCFVLDKDVITYSYTTGFMNSLFSGNFAK 102
XF0262      GMVYGGIAGAIGGAIGFVLDKDVITYSYTNGFMSSIFNGTFAK 102
2_J1a12     GMVYGGVGLVGGGLISGFIYGKERAYAYSDFMGSLFNGTFAK 103
          *****:. *  :**  .  *:  . *  *  *:  **  .*:*.*.*:.*

```

Figura 15. Comparação das sequências de colicinas V da cepa 9a5c e J1a12. Alinhamento múltiplo das sequências de aminoácidos codificadas pelas CDS XF0262, XF0263 e XF0264 anotadas no genoma da cepa 9a5c e das duas CDS ortólogas identificadas na sequência do genoma de J1a12. As sequências foram alinhadas utilizando-se o programa CLUSTAL W. (*) indica resíduos idênticos, (:) indica substituição conservativa e (.) indica substituição semi-conservativa.

Figura 16. Resultado da busca no GenBank por sequências similares ao contig 374. A sequência de nucleotídeos do contig 374 foi utilizada para interrogar o banco de dados nr/nt do GenBank utilizando-se a ferramenta BLASTn. A figura é uma montagem do resultado obtido, destacando os melhores hits e algumas das características do alinhamento, como a cobertura e a identidade máxima observada.

Contig 374

CDS	Produto predito
1	ATPase provides energy for both assembly of type IV secretion complex and secretion of T-DNA complex (VirB4)
2	Minor pilin of type IV secretion complex (VirB5)
3	Integral inner membrane protein of type IV secretion complex (VirB6)
4	Inner membrane protein forms channel for type IV secretion of T-DNA complex (VirB8)
5	Forms the bulk of type IV secretion complex that spans outer membrane and periplasm (VirB9)
6	Inner membrane protein forms channel for type IV secretion of T-DNA complex (VirB10)
7	ATPase provides energy for both assembly of type IV secretion complex and secretion of T-DNA complex (VirB11)
8	Type IV secretion system protein VirD4
9	hypothetical protein
10	hypothetical protein
11	hypothetical protein
12	hypothetical protein
13	IncQ plasmid conjugative transfer DNA nicking endonuclease TraR (pTi VirD2 homolog)
14	putative virulence protein
15	Transposase
16	Programmed cell death toxin PemK
17	Programmed cell death antitoxin PemI
18	Zeta toxin family protein
19	hypothetical protein
20	mobilisation protein
21	hypothetical protein
22	hypothetical protein
23	transfer origin protein, TraL
24	Resolvase, N-terminal domain
25	putative replication protein A

Figura 17. Anotação preliminar do contig 374. As CDS, com a orientação de transcrição indicadas pelo sentido das setas, foram preditas a partir da análise da sequência do contig 374 com as ferramentas BLASTx e PSI-BLAST associadas a anotação automática realizada no servidor RAST. O produto predito para todas as CDS esquematizadas está indicado.

5. Discussão Final

O sequenciamento completo e anotação do genoma de diversas cepas de *Xylella fastidiosa* revelaram uma marcante similaridade entre seus respectivos repertórios de genes, ainda que estas cepas sejam específicas para os diferentes hospedeiros vegetais que colonizam (Van Sluys *et al.*, 2003; Chatterjee *et al.*, 2008a; de Mello Varani *et al.*, 2009). Os fatores que respondem por esta especificidade ao hospedeiro ainda não foram elucidados. Por outro lado, nos últimos anos, abordagens genômicas e de genética molecular possibilitaram a identificação e o estudo funcional de potenciais determinantes de virulência e patogenicidade comuns às distintas cepas de *X. fastidiosa*. Aspectos da fisiologia e da regulação gênica deste patógeno também têm sido bastante estudados, revelando as similitudes e as diferenças com outras bactérias patogênicas.

Por exemplo, o genoma de *X. fastidiosa* não possui genes do sistema de secreção tipo III que é conservado na maioria dos fitopatógenos Gram-negativos e que é responsável pela translocação de fatores de virulência (efetores) diretamente para célula vegetal. Por outro lado, igualmente a certos fitopatógenos, o genoma de *X. fastidiosa* codifica outros sistemas de secreção de proteínas, por exemplo o sistema de secreção do tipo II, que possivelmente é responsável pela exportação de enzimas necessárias para degradação da parede celular da planta, como poligalacturonases e endoglucanases. A virulência de *X. fastidiosa* também está associada a sua motilidade *twitching* dependente de *pilus* do tipo IV e também de adesinas fimbriais e não-fimbriais, fundamentais para adesão intercelular e à superfície durante a colonização de seus hospedeiros (planta e inseto), garantindo a formação do biofilme, que assim como em outros patógenos, é um importante determinante de patogenicidade. Além disso, também como outros fitopatógenos, como os do gênero *Xanthomonas*, *X. fastidiosa* produz polissacarídeo extracelular (EPS) que parece ser importante na maturação do biofilme (Van Sluys *et al.*, 2002; Van Sluys *et al.*, 2003; Guilhabert e Kirkpatrick, 2005; Meng *et al.*, 2005; Reddy *et al.*, 2007; Roper *et al.*, 2007a; Roper *et al.*, 2007b; Chatterjee *et al.*, 2008a; Holtsmark *et al.*, 2008; Killiny e Almeida, 2008; Newman *et al.*, 2008; Rudrappa *et al.*, 2008; de Mello Varani *et al.*, 2009; Kay e Bonas, 2009; Killiny e Almeida, 2009b).

O modelo atual de virulência e patogenicidade de *X. fastidiosa* contempla todos estes fatores, e postula que a colonização do hospedeiro vegetal e sua transmissão pelo inseto vetor são processos mutuamente exclusivos, controlados em grande parte pela extensão da colonização dos vasos do xilema. Isto é, o comportamento da bactéria na planta é coordenado, entre uma colonização sistêmica inicial e o aumento localizado da população bacteriana nos vasos do xilema, sendo esta uma situação mais favorável a sua aquisição pelo inseto vetor. Este comportamento é controlado pelo acúmulo de DSF (*Diffusible*

Signaling Factor) uma molécula sinalizadora cuja síntese acompanha o aumento da densidade celular, como ocorre em diversos outros fitopatógenos (da Silva *et al.*, 2001; Scarpari *et al.*, 2003; Wang *et al.*, 2004; Colnaghi Simionato *et al.*, 2007; Chatterjee *et al.*, 2008a; Chatterjee *et al.*, 2008c; Dow, 2008; Newman *et al.*, 2008). Porém, considerando a complexidade, a especificidade e seletividade da interação de *X. fastidiosa* com seus distintos hospedeiros vegetais e com o inseto vetor, presume-se que fatores e mecanismos adicionais estejam envolvidos nestes processos (Chatterjee *et al.*, 2008a).

Neste trabalho, nós mostramos que a expressão gênica de *X. fastidiosa* 9a5c é modulada pela exposição à dose subletal do antibiótico estreptomicina, de modo similar a outras bactérias (Ng *et al.*, 2003; Qiu *et al.*, 2005). O conjunto de genes modulados difere do conjunto de genes modulados pela exposição de *X. fastidiosa* à dose subletal do peptídeo antimicrobiano gomesina (Fogaça *et al.*, 2010) e inclui CDS relacionadas à resposta de choque térmico além de CDS associadas ao processo de tradução. Interessantemente, a exposição à estreptomicina também induz a expressão da CDS que codifica poligalacturonase, que é um fator de virulência em algumas cepas de *X. fastidiosa*. O gene de poligalacturonase de *X. fastidiosa* 9a5c, uma cepa virulenta originalmente isolada de laranjeiras com sintomas de clorose variegada dos citros, apresenta uma mutação *frameshift*. Portanto, a hipótese atual é que a falta de poligalacturonase explicaria o fenótipo menos agressivo desta cepa comparativamente à cepa Temecula que causa a doença de Pierce das videiras (Van Sluys *et al.*, 2003). Por outro lado, o tratamento com estreptomicina promoveu a modulação negativa de CDS relacionadas à formação e manutenção de biofilme ao contrário do observado quando estas bactérias foram submetidas ao tratamento com gomesina (Fogaça *et al.*, 2010). O conjunto destas observações sugere que a exposição à dose subletal de estreptomicina possa promover um fenótipo de maior virulência, contrariamente ao efeito previamente observado com a gomesina. Neste contexto, acreditamos que seria importante avaliar o efeito destes antibióticos em cepas virulentas de *X. fastidiosa* que não possuam a mutação do gene da poligalacturonase.

A comparação do perfil de expressão gênica de *X. fastidiosa* exibido em resposta ao tratamento com estreptomicina com perfis de expressão observados em outras condições experimentais, se por um lado demonstrou a especificidade da resposta a estreptomicina, por outro lado apontou diversas CDS de função desconhecida cujos respectivos perfis de expressão avalizariam uma caracterização mais detalhada. Como parte do trabalho desenvolvido durante o Doutorado, pretendíamos pesquisar possíveis funções de algumas CDS selecionadas com este critério. A abordagem escolhida foi a de identificar potenciais parceiros moleculares dos produtos codificados por estas CDS utilizando o sistema de duplo-híbrido em bactérias como descrito por (Karimova *et al.*, 1998; Karimova *et al.*, 2005). Etapas importantes para implantação desta metodologia foram cumpridas, mas infelizmente

não conseguimos obter uma biblioteca genômica adequada para buscarmos prováveis interações com CDS selecionadas. Este projeto será continuado em nosso laboratório, associado a outras abordagens para caracterização da função de CDS selecionadas. Vale ressaltar que a dificuldade na manipulação genética de *X. fastidiosa*, em particular a da cepa 9a5c, é um impecilho para utilização de abordagens genéticas para caracterização funcional de proteínas.

Finalmente, os resultados da análise do pirosequenciamento do genoma da cepa J1a12 de *Xylella fastidiosa*, que exibe fenótipo menos virulento em citros e tabaco em relação à cepa 9a5c, são extremamente promissores para explicar estas diferenças fenotípicas. É importante mencionar que o sequenciamento que realizamos confirmou diferenças entre os genomas destas duas cepas e que haviam sido anteriormente observadas utilizando-se microarranjos de DNA (Koide *et al.*, 2004b). Nossos resultados indicam outras diferenças importantes que, possivelmente, serão úteis para explicar o fenótipo menos virulento da cepa J1a12 bem como sua suscetibilidade à transformação por DNA, a qual tem possibilitado a manipulação genética desta cepa. Entre as diferenças adicionais identificadas pelo pirosequenciamento está a observação de que a cepa J1a12 possui poliformismos nos genes que codificam para sistema de restrição e modificação de DNA. Além disso, as sequências já obtidas indicam a presença de um plasmídeo adicional na cepa J1a12, além do já conhecido pXF51.

6. Referências Bibliográficas

1. **Ahmadian, A., Ehn, M. e Hober, S.** 2006. Pyrosequencing: history, biochemistry and future. *Clin Chim Acta* 363:83-94.
2. **Allen, T. D., Watkins, T., Lindahl, L. e Zengel, J. M.** 2004. Regulation of ribosomal protein synthesis in *Vibrio cholerae*. *J Bacteriol* 186:5933-7.
3. **Almeida, L. G., Paixao, R., Souza, R. C., Costa, G. C., Barrientos, F. J., Santos, M. T., Almeida, D. F. e Vasconcelos, A. T.** 2004. A System for Automated Bacterial (genome) Integrated Annotation--SABIA. *Bioinformatics* 20:2832-3.
4. **Almeida, R. P., Nascimento, F. E., Chau, J., Prado, S. S., Tsai, C. W., Lopes, S. A. e Lopes, J. R.** 2008. Genetic structure and biology of *Xylella fastidiosa* strains causing disease in citrus and coffee in Brazil. *Appl Environ Microbiol* 74:3690-701.
5. **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 the Entomological Society of America* 98:775-786.
6. **Altschul, S. F., Gish, W., Miller, W., Myers, E. W. e Lipman, D. J.** 1990. Basic local alignment search tool. *J Mol Biol* 215:403-10.
7. **Altschul, S. F., Madden, T. L., Schaffer, 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 Res* 25:3389-402.
8. **Alvarez-Martinez, C. E. e Christie, P. J.** 2009. Biological diversity of prokaryotic type IV secretion systems. *Microbiol Mol Biol Rev* 73:775-808.
9. **Alves, E., Kitajima, E. W. e Leite, B.** 2003. Interaction of *Xylella fastidiosa* with different cultivars of *Nicotiana tabacum*: a comparison of colonization patterns. *Journal of Phytopathology-Phytopathologische Zeitschrift* 151:500-506.
10. **Ausubel, F. M. e Frederick, M.** 1995. Current protocols in molecular biology. John Wiley & Sons, New York.
11. **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* 9:75.
12. **Bader, M. W., Navarre, W. W., Shiau, W., Nikaido, H., Frye, J. G., McClelland, M., Fang, F. C. e Miller, S. I.** 2003. Regulation of *Salmonella typhimurium* virulence gene expression by cationic antimicrobial peptides. *Mol Microbiol* 50:219-30.
13. **Baichoo, N., Wang, T., Ye, R. e Helmann, J. D.** 2002. Global analysis of the *Bacillus subtilis* Fur regulon and the iron starvation stimulon. *Mol Microbiol* 45:1613-29.
14. **Barber, C. E., Tang, J. L., Feng, J. X., Pan, M. Q., Wilson, T. J., Slater, H., Dow, J. M., Williams, P. e Daniels, M. J.** 1997. A novel regulatory system required for pathogenicity of *Xanthomonas campestris* is mediated by a small diffusible signal molecule. *Mol Microbiol* 24:555-66.
15. **Barbosa, R. L. e Benedetti, C. E.** 2007. BigR, a Transcriptional Repressor from Plant-Associated Bacteria, Regulates an Operon Implicated in Biofilm Growth. *J Bacteriol*.
16. **Bertone, P., Stolc, V., Royce, T. E., Rozowsky, J. S., Urban, A. E., Zhu, X., Rinn, J. L., Tongprasit, W., Samanta, M., Weissman, S., Gerstein, M. e Snyder, M.** 2004.

- Global identification of human transcribed sequences with genome tiling arrays. *Science* 306:2242-6.
17. **Betts, J. C., McLaren, A., Lennon, M. G., Kelly, F. M., Lukey, P. T., Blakemore, S. J. e Duncan, K.** 2003. Signature gene expression profiles discriminate between isoniazid-, thiolactomycin-, and triclosan-treated *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother* 47:2903-13.
 18. **Bhattacharyya, A., Stilwagen, S., Ivanova, N., D'Souza, M., Bernal, A. et al.** 2002a. Whole-genome comparative analysis of three phytopathogenic *Xylella fastidiosa* strains. *Proc Natl Acad Sci U S A* 99:12403-8.
 19. **Bhattacharyya, A., Stilwagen, S., Reznik, G., Feil, H., Feil, W. S. et al.** 2002b. Draft sequencing and comparative genomics of *Xylella fastidiosa* strains reveal novel biological insights. *Genome Res* 12:1556-63.
 20. **Boni, I. V., Artamonova, V. S., Tzareva, N. V. e Dreyfus, M.** 2001. Non-canonical mechanism for translational control in bacteria: synthesis of ribosomal protein S1. *Embo J* 20:4222-32.
 21. **Brazas, M. D. e Hancock, R. E.** 2005. Using microarray gene signatures to elucidate mechanisms of antibiotic action and resistance. *Drug Discov Today* 10:1245-52.
 22. **Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P. et al.** 2001. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet* 29:365-71.
 23. **Brazma, A. e Vilo, J.** 2000. Gene expression data analysis. *FEBS Lett* 480:17-24.
 24. **Burrack, L. S. e Higgins, D. E.** 2007. Genomic approaches to understanding bacterial virulence. *Curr Opin Microbiol* 10:4-9.
 25. **Caldana, C.** 2002. Avaliação da expressão gênica de *Xylella fastidiosa* em condições de adesão. Dissertação de Mestrado. Instituto de Biologia, Universidade Estadual de Campinas, Campinas. 79 p.
 26. **Carter, A. P., Clemons, W. M., Brodersen, D. E., Morgan-Warren, R. J., Wimberly, B. T. e Ramakrishnan, V.** 2000. Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* 407:340-8.
 27. **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.
 28. **Chatterjee, S., Almeida, R. P. e Lindow, S.** 2008a. Living in two worlds: the plant and insect lifestyles of *Xylella fastidiosa*. *Annu Rev Phytopathol* 46:243-71.
 29. **Chatterjee, S., Newman, K. L. e Lindow, S. E.** 2008b. Cell-to-cell signaling in *Xylella fastidiosa* suppresses movement and xylem vessel colonization in grape. *Molecular Plant-Microbe Interactions* 21:1309-1315.
 30. **Chatterjee, S., Wistrom, C. e Lindow, S. E.** 2008c. A cell-cell signaling sensor is required for virulence and insect transmission of *Xylella fastidiosa*. *Proc Natl Acad Sci U S A* 105:2670-5.
 31. **Chen, J., Banks, D., Jarret, R. L., Chang, C. J. e Smith, B. J.** 2000. Use of 16S rDNA sequences as signature characters to identify *Xylella fastidiosa*. *Curr Microbiol* 40:29-33.
 32. **Chen, J., Chang, C. J., Jarret, R. L. e Gawel, N.** 1992. Genetic variation among

- Xylella fastidiosa* Strains. *Phytopathology* 82:973-977.
33. **Chen, J., Han, S., Civerolo, E., Stenger, D. C. e Van Sluys, M.** 2007. Two whole genome sequences of *Xylella fastidiosa* almond leaf scorch strains. *Phytopathology* 97:S21-S22.
 34. **Coletta-Filho, H. D. e Machado, M. A.** 2003. Geographical Genetic Structure of *Xylella fastidiosa* from Citrus in Sao Paulo State, Brazil. *Phytopathology* 93:28-34.
 35. **Colnaghi Simionato, A. V., da Silva, D. S., Lambais, M. R. e Carrilho, E.** 2007. Characterization of a putative *Xylella fastidiosa* diffusible signal factor by HRGC-EI-MS. *J Mass Spectrom* 42:490-6.
 36. **Conway, T. e Schoolnik, G. K.** 2003. Microarray expression profiling: capturing a genome-wide portrait of the transcriptome. *Mol Microbiol* 47:879-89.
 37. **Costa de Oliveira, R., Yanai, G. M., Muto, N. H., Leite, D. B., de Souza, A. A., Coletta Filho, H. D., Machado, M. A. e Nunes, L. R.** 2002. Competitive hybridization on spotted microarrays as a tool to conduct comparative genomic analyses of *Xylella fastidiosa* strains. *FEMS Microbiol Lett* 216:15-21.
 38. **Cussiol, J. R., Alves, S. V., de Oliveira, M. A. e Netto, L. E.** 2003. Organic hydroperoxide resistance gene encodes a thiol-dependent peroxidase. *J Biol Chem* 278:11570-8.
 39. **da Costa, P. I., Franco, C. F., Miranda, V. S., Teixeira, D. C. e Hartung, J. S.** 2000. Strains of *Xylella fastidiosa* rapidly distinguished by arbitrarily primed-PCR. *Curr Microbiol* 40:279-82.
 40. **da Silva, F. R., Vettore, A. L., Kemper, E. L., Leite, A. e Arruda, P.** 2001. Fastidian gum: the *Xylella fastidiosa* exopolysaccharide possibly involved in bacterial pathogenicity. *Fems Microbiology Letters* 203:165-171.
 41. **da Silva Neto, J. F., Koide, T., Abe, C. M., Gomes, S. L. e Marques, M. V.** 2008. Role of sigma54 in the regulation of genes involved in type I and type IV pili biogenesis in *Xylella fastidiosa*. *Arch Microbiol* 189:249-61.
 42. **da Silva Neto, J. F., Koide, T., Gomes, S. L. e Marques, M. V.** 2002. Site-directed gene disruption in *Xylella fastidiosa*. *FEMS Microbiol Lett* 210:105-10.
 43. **da Silva Neto, J. F., Koide, T., Gomes, S. L. e Marques, M. V.** 2007. The single extracytoplasmic-function sigma factor of *Xylella fastidiosa* is involved in the heat shock response and presents an unusual regulatory mechanism. *J Bacteriol* 189:551-60.
 44. **da Silva, V. S., Shida, C. S., Rodrigues, F. B., Ribeiro, D. C., de Souza, A. A., Coletta-Filho, H. D., Machado, M. A., Nunes, L. R. e de Oliveira, R. C.** 2007. Comparative genomic characterization of citrus-associated *Xylella fastidiosa* strains. *BMC Genomics* 8:474.
 45. **Davis, M. J., French, W. J. e Schaad, N. W.** 1981. Axenic culture of the bacteria associated with phony disease of peach and plum leaf scald. *Current Microbiology* 6:309-314.
 46. **De La Fuente, L., Burr, T. J. e Hoch, H. C.** 2007. Mutations in type I and type IV pilus biosynthetic genes affect twitching motility rates in *Xylella fastidiosa*. *J Bacteriol* 189:7507-10.
 47. **De La Fuente, L., Burr, T. J. e Hoch, H. C.** 2008. Autoaggregation of *Xylella fastidiosa* cells is influenced by type I and type IV pili. *Appl Environ Microbiol*.

48. **de Mello Varani, A., de Lima, W. C., Moreira, L. M., Oliveira, M., Souza, R. C., Civerolo, E., Vasconcelos, A. T. R. e Van Sluys, M. A.** 2009. Common Genes and Genomic Breaks: A Detailed Case Study of the *Xylella fastidiosa* Genome Backbone and Evolutionary Insights, p. 330. In R. W. Jackson (ed.), Plant Pathogenic Bacteria: Genomics and Molecular Biology. Caister Academic Press, Norwich.
49. **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.
50. **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. Mol Plant Microbe Interact 16:867-75.
51. **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 Microbiol Lett 237:341-53.
52. **de Souza, A. A., Takita, M. A., Pereira, E. O., Coletta-Filho, H. D. e Machado, M. A.** 2005. Expression of pathogenicity-related genes of *Xylella fastidiosa* in vitro and in planta. Curr Microbiol 50:223-8.
53. **DeLisa, M. P., Wu, C. F., Wang, L., Valdes, J. J. e Bentley, W. E.** 2001. DNA microarray-based identification of genes controlled by autoinducer 2-stimulated quorum sensing in *Escherichia coli*. Journal of Bacteriology 183:5239-5247.
54. **Deslandes, V., Nash, J. H., Harel, J., Coulton, J. W. e Jacques, M.** 2007. Transcriptional profiling of *Actinobacillus pleuropneumoniae* under iron-restricted conditions. BMC Genomics 8:72.
55. **Do, T. Q., Schultz, J. R. e Clarke, C. F.** 1996. Enhanced sensitivity of ubiquinone-deficient mutants of *Saccharomyces cerevisiae* to products of autoxidized polyunsaturated fatty acids. Proc Natl Acad Sci U S A 93:7534-9.
56. **Doddapaneni, H., Yao, J., Lin, H., Walker, M. A. e Civerolo, E. L.** 2006. Analysis of the genome-wide variations among multiple strains of the plant pathogenic bacterium *Xylella fastidiosa*. BMC Genomics 7:225.
57. **Dow, M.** 2008. Diversification of the function of cell-to-cell signaling in regulation of virulence within plant pathogenic xanthomonads. Sci Signal 1:pe23.
58. **Droege, M. e Hill, B.** 2008. The Genome Sequencer FLX System--longer reads, more applications, straight forward bioinformatics and more complete data sets. J Biotechnol 136:3-10.
59. **Edgar, R. e Barrett, T.** 2006. NCBI GEO standards and services for microarray data. Nat Biotechnol 24:1471-2.
60. **Eisen, M. B., Spellman, P. T., Brown, P. O. e Botstein, D.** 1998. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 95:14863-8.
61. **Ernst, R. K., Guina, T. e Miller, S. I.** 2001. *Salmonella typhimurium* outer membrane remodeling: role in resistance to host innate immunity. Microbes Infect 3:1327-34.
62. **Evers, S., Di Padova, K., Meyer, M., Langen, H., Fountoulakis, M., Keck, W. e Gray, C. P.** 2001. Mechanism-related changes in the gene transcription and protein

- synthesis patterns of *Haemophilus influenzae* after treatment with transcriptional and translational inhibitors. *Proteomics* 1:522-44.
63. **Fitzgerald, J. R., Sturdevant, D. E., Mackie, S. M., Gill, S. R. e Musser, J. M.** 2001. Evolutionary genomics of *Staphylococcus aureus*: insights into the origin of methicillin-resistant strains and the toxic shock syndrome epidemic. *Proc Natl Acad Sci U S A* 98:8821-6.
 64. **Fogaça, A. C., Zaini, P. A., Wulff, N. A., da Silva, P. I. P., Fázio, M. A., Miranda, A., Daffre, S. e da Silva, A. M.** 2010. Effects of the antimicrobial peptide gomesin on the global gene expression profile, virulence, and biofilm formation of *Xylella fastidiosa*. submetido à publicação.
 65. **Francis, M., Civerolo, E. L. e Bruening, G.** 2008. Improved bioassay of *Xylella fastidiosa* strains using *Nicotiana tabacum* cultivar SR1. *Plant Disease* 92:14-20.
 66. **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. *J Exp Bot* 58:4037-46.
 67. **Goff, S. A. e Goldberg, A. L.** 1985. Production of abnormal proteins in *E. coli* stimulates transcription of lon and other heat shock genes. *Cell* 41:587-95.
 68. **Goh, E. B., Yim, G., Tsui, W., McClure, J., Surette, M. G. e Davies, J.** 2002. Transcriptional modulation of bacterial gene expression by subinhibitory concentrations of antibiotics. *Proc Natl Acad Sci U S A* 99:17025-30.
 69. **Groisman, E. A., Chiao, E., Lipps, C. J. e Heffron, F.** 1989. *Salmonella typhimurium* phoP virulence gene is a transcriptional regulator. *Proc Natl Acad Sci U S A* 86:7077-81.
 70. **Gromadski, K. B. e Rodnina, M. V.** 2004a. Kinetic determinants of high-fidelity tRNA discrimination on the ribosome. *Mol Cell* 13:191-200.
 71. **Gromadski, K. B. e Rodnina, M. V.** 2004b. Streptomycin interferes with conformational coupling between codon recognition and GTPase activation on the ribosome. *Nat Struct Mol Biol* 11:316-22.
 72. **Guilhbert, M. R. e Kirkpatrick, B. C.** 2005. Identification of *Xylella fastidiosa* antivirulence genes: hemagglutinin adhesins contribute a biofilm maturation to *X. fastidiosa* and colonization and attenuate virulence. *Mol Plant Microbe Interact* 18:856-68.
 73. **Hansen, E. H., Schembri, M. A., Klemm, P., Schafer, T., Molin, S. e Gram, L.** 2004. Elucidation of the antibacterial mechanism of the *Curvularia haloperoxidase* system by DNA microarray profiling. *Appl Environ Microbiol* 70:1749-57.
 74. **Havarstein, L. S., Holo, H. e Nes, I. F.** 1994. The leader peptide of colicin V shares consensus sequences with leader peptides that are common among peptide bacteriocins produced by gram-positive bacteria. *Microbiology* 140 (Pt 9):2383-9.
 75. **Highlander, S. K., Hulten, K. G., Qin, X., Jiang, H., Yerrapragada, S. et al.** 2007. Subtle genetic changes enhance virulence of methicillin resistant and sensitive *Staphylococcus aureus*. *BMC Microbiol* 7:99.
 76. **Hoffman, L. R., D'Argenio, D. A., MacCoss, M. J., Zhang, Z., Jones, R. A. e Miller, S. I.** 2005. Aminoglycoside antibiotics induce bacterial biofilm formation. *Nature* 436:1171-5.
 77. **Hoheisel, J. D.** 2006. Microarray technology: beyond transcript profiling and

- genotype analysis. *Nat Rev Genet* 7:200-10.
78. **Holberger, L. E. e Hayes, C. S.** 2009. Ribosomal protein S12 and aminoglycoside antibiotics modulate A-site mRNA cleavage and transfer-messenger RNA activity in *Escherichia coli*. *J Biol Chem* 284:32188-200.
 79. **Holloway, A. J., van Laar, R. K., Tothill, R. W. e Bowtell, D. D.** 2002. Options available--from start to finish--for obtaining data from DNA microarrays II. *Nat Genet* 32 Suppl:481-9.
 80. **Holt, K. E., Parkhill, J., Mazzoni, C. J., Roumagnac, P., Weill, F. X., Goodhead, I., Rance, R., Baker, S., Maskell, D. J., Wain, J., Dolecek, C., Achtman, M. e Dougan, G.** 2008. High-throughput sequencing provides insights into genome variation and evolution in *Salmonella typhi*. *Nat Genet* 40:987-93.
 81. **Holtmark, I., Eijsink, V. G. H. e Brurberg, M. B.** 2008. Bacteriocins from plant pathogenic bacteria. *FEMS Microbiology Letters* 280:1-7.
 82. **Hong, R. W., Shchepetov, M., Weiser, J. N. e Axelsen, P. H.** 2003. Transcriptional profile of the *Escherichia coli* response to the antimicrobial insect peptide cecropin A. *Antimicrob Agents Chemother* 47:1-6.
 83. **Hopkins, D. L.** 1989. *Xylella fastidiosa* - Xylem-Limited Bacterial Pathogen of Plants. *Annual Review of Phytopathology* 27:271-290.
 84. **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.
 85. **Horwich, A. L., Fenton, W. A., Chapman, E. e Farr, G. W.** 2007. Two families of chaperonin: physiology and mechanism. *Annu Rev Cell Dev Biol* 23:115-45.
 86. **Jones, A., Georg, M., Maudsdotter, L. e Jonsson, A. B.** 2009. Endotoxin, capsule, and bacterial attachment contribute to *Neisseria meningitidis* resistance to the human antimicrobial peptide LL-37. *J Bacteriol* 191:3861-8.
 87. **Jones, J. D. e Dangl, J. L.** 2006. The plant immune system. *Nature* 444:323-9.
 88. **Karimova, G., Dautin, N. e Ladant, D.** 2005. Interaction network among *Escherichia coli* membrane proteins involved in cell division as revealed by bacterial two-hybrid analysis. *J Bacteriol* 187:2233-43.
 89. **Karimova, G., Pidoux, J., Ullmann, A. e Ladant, D.** 1998. A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc Natl Acad Sci U S A* 95:5752-6.
 90. **Kaushal, D., Schroeder, B. G., Tyagi, S., Yoshimatsu, T., Scott, C., Ko, C., Carpenter, L., Mehrotra, J., Manabe, Y. C., Fleischmann, R. D. e Bishai, W. R.** 2002. Reduced immunopathology and mortality despite tissue persistence in a *Mycobacterium tuberculosis* mutant lacking alternative sigma factor, SigH. *Proc Natl Acad Sci U S A* 99:8330-5.
 91. **Kay, S. e Bonas, U.** 2009. How *Xanthomonas* type III effectors manipulate the host plant. *Curr Opin Microbiol* 12:37-43.
 92. **Killiny, N. e Almeida, R.** 2008. Role of fimbrial and afimbrial adhesins and gum production on *Xylella fastidiosa* insect transmission. *Phytopathology* 98:S81-S81.
 93. **Killiny, N. e Almeida, R. P.** 2009a. Host structural carbohydrate induces vector transmission of a bacterial plant pathogen. *Proc Natl Acad Sci U S A* 106:22416-20.
 94. **Killiny, N. e Almeida, R. P.** 2009b. *Xylella fastidiosa* afimbrial adhesins mediate cell

- transmission to plants by leafhopper vectors. *Appl Environ Microbiol* 75:521-8.
95. **Koide, T.** 2006. Análise da global da expressão gênica de *Xylella fastidiosa* submetida a estresses ambientais. Tese de Doutorado. Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, São Paulo. 179 p.
 96. **Koide, T., da Silva Neto, J. F., Gomes, S. L. e Marques, M. V.** 2004a. Insertional transposon mutagenesis in the *Xylella fastidiosa* Citrus Variegated Chlorosis strain with transposome. *Curr Microbiol* 48:247-50.
 97. **Koide, T., Salem-Izacc, S. M., Gomes, S. L. e Vencio, R. Z.** 2006a. SpotWhatR: a user-friendly microarray data analysis system. *Genet Mol Res* 5:93-107.
 98. **Koide, T., Vencio, R. Z. e Gomes, S. L.** 2006b. Global gene expression analysis of the heat shock response in the phytopathogen *Xylella fastidiosa*. *J Bacteriol* 188:5821-30.
 99. **Koide, T., Zaini, P. A., Moreira, L. M., Vencio, R. Z., Matsukuma, A. Y., Durham, A. M., Teixeira, D. C., El-Dorry, H., Monteiro, P. B., da Silva, A. C., Verjovski-Almeida, S., da Silva, A. M. e Gomes, S. L.** 2004b. DNA microarray-based genome comparison of a pathogenic and a nonpathogenic strain of *Xylella fastidiosa* delineates genes important for bacterial virulence. *J Bacteriol* 186:5442-9.
 100. **Laub, M. T., McAdams, H. H., Feldblyum, T., Fraser, C. M. e Shapiro, L.** 2000. Global analysis of the genetic network controlling a bacterial cell cycle. *Science* 290:2144-8.
 101. **Lerat, E., Daubin, V., Ochman, H. e Moran, N. A.** 2005. Evolutionary origins of genomic repertoires in bacteria. *PLoS Biol* 3:e130.
 102. **Li, W. B., Zreik, L., Fernandes, N. G., Miranda, V. S., Teixeira, D. C., Ayres, A. J., Garnier, M. e Bov, J. M.** 1999. A triply cloned strain of *Xylella fastidiosa* multiplies and induces symptoms of citrus variegated chlorosis in sweet orange. *Curr Microbiol* 39:106-8.
 103. **Li, Y., Hao, G., Galvani, C. D., Meng, Y., De La Fuente, L., Hoch, H. C. e Burr, T. J.** 2007. Type I and type IV pili of *Xylella fastidiosa* affect twitching motility, biofilm formation and cell-cell aggregation. *Microbiology* 153:719-26.
 104. **Lin, H., Doddapaneni, H., Takahashi, Y. e Walker, M. A.** 2007. Comparative analysis of ESTs involved in grape responses to *Xylella fastidiosa* infection. *BMC Plant Biol* 7:8.
 105. **Linares, J. F., Gustafsson, I., Baquero, F. e Martinez, J. L.** 2006. Antibiotics as intermicrobial signaling agents instead of weapons. *Proc Natl Acad Sci U S A* 103:19484-9.
 106. **Lioy, V. S., Rey, O., Balsa, D., Pellicer, T. e Alonso, J. C.** 2010. A toxin-antitoxin module as a target for antimicrobial development. *Plasmid* 63:31-9.
 107. **Livak, K. J. e Schmittgen, T. D.** 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25:402-8.
 108. **Lockhart, D. J. e Winzeler, E. A.** 2000. Genomics, gene expression and DNA arrays. *Nature* 405:827-36.
 109. **Lopes, S. A., Ribeiro, D. M., Roberto, P. G., Franca, S. C. e Santos, J. M.** 2000. *Nicotiana tabacum* as an experimental host for the study of plant-*Xylella fastidiosa* interactions. *Plant Disease* 84:827-830.

110. **Lucchini, S., Thompson, A. e Hinton, J. C.** 2001. Microarrays for microbiologists. *Microbiology* 147:1403-14.
111. **Lund, P. A.** 2009. Multiple chaperonins in bacteria--why so many? *FEMS Microbiol Rev* 33:785-800.
112. **Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S. et al.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437:376-80.
113. **Maria Fedatto, L., Silva-Stenico, M. E., Etchegaray, A., Pacheco, F. T., Rodrigues, J. L. e Tsai, S. M.** 2006. Detection and characterization of protease secreted by the plant pathogen *Xylella fastidiosa*. *Microbiol Res* 161:263-72.
114. **Marques, M. V. e da Silva, A. M.** 2004. Transcriptoma, p. 119-137. *In* L. Mir (ed.), *Genômica*. Atheneu, São Paulo.
115. **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.
116. **Maurer, A. P., Mehlitz, A., Mollenkopf, H. J. e Meyer, T. F.** 2007. Gene Expression Profiles of *Chlamydophila pneumoniae* during the Developmental Cycle and Iron Depletion-Mediated Persistence. *PLoS Pathog* 3:e83.
117. **McElrone, A. J., Jackson, S. e Habdas, P.** 2008. Hydraulic disruption and passive migration by a bacterial pathogen in oak tree xylem. *J Exp Bot* 59:2649-57.
118. **McElrone, A. J., Sherald, J. L. e Forseth, I. N.** 2001. Effects of water stress on symptomatology and growth of *Parthenocissus quinquefolia* infected by *Xylella fastidiosa*. *Plant Disease* 85:1160-1164.
119. **McElrone, A. J., Sherald, J. L. e Forseth, I. N.** 2003. Interactive effects of water stress and xylem-limited bacterial infection on the water relations of a host vine. *Journal of Experimental Botany* 54:419-430.
120. **McManus, P. S., Stockwell, V. O., Sundin, G. W. e Jones, A. L.** 2002. Antibiotic use in plant agriculture. *Annu Rev Phytopathol* 40:443-65.
121. **Mehta, A., Leite, R. P., Jr. e Rosato, Y. B.** 2001. Assessment of the genetic diversity of *Xylella fastidiosa* isolated from citrus in Brazil by PCR-RFLP of the 16S rDNA and 16S-23S intergenic spacer and rep-PCR fingerprinting. *Antonie Van Leeuwenhoek* 79:53-9.
122. **Meidanis, J., Braga, M. D. e Verjovski-Almeida, S.** 2002. Whole-genome analysis of transporters in the plant pathogen *Xylella fastidiosa*. *Microbiol Mol Biol Rev* 66:272-99.
123. **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. *J Bacteriol* 187:5560-7.
124. **Michel, A., Agerer, F., Hauck, C. R., Herrmann, M., Ullrich, J., Hacker, J. e Ohlsen, K.** 2006. Global regulatory impact of ClpP protease of *Staphylococcus aureus* on regulons involved in virulence, oxidative stress response, autolysis, and DNA repair. *J Bacteriol* 188:5783-96.
125. **Mingeot-Leclercq, M. P., Glupczynski, Y. e Tulkens, P. M.** 1999. Aminoglycosides: activity and resistance. *Antimicrob Agents Chemother* 43:727-37.
126. **Mockler, T. C., Chan, S., Sundaresan, A., Chen, H., Jacobsen, S. E. e Ecker, J. R.** 2005. Applications of DNA tiling arrays for whole-genome analysis. *Genomics* 85:1-

- 15.
127. **Mogk, A., Volker, A., Engelmann, S., Hecker, M., Schumann, W. e Volker, U.** 1998. Nonnative proteins induce expression of the *Bacillus subtilis* CIRCE regulon. *J Bacteriol* 180:2895-900.
128. **Monteiro, P. B., Renaudin, J., Jagoueix-Eveillard, S., Ayres, A. J., Garnier, M. e Bove, J. M.** 2001a. *Catharanthus roseus*, an experimental host plant for the citrus strain of *Xylella fastidiosa*. *Plant Disease* 85:246-251.
129. **Monteiro, P. B., Teixeira, D. C., Palma, R. R., Garnier, M., Bove, J. M. e Renaudin, J.** 2001b. Stable transformation of the *Xylella fastidiosa* citrus variegated chlorosis strain with oriC plasmids. *Appl Environ Microbiol* 67:2263-9.
130. **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.
131. **Munro, S. e Pelham, H.** 1985. What turns on heat shock genes? *Nature* 317:477-8.
132. **Neves, E. M., Dayoub, M., Dragone, D. S. e Neves, M. F.** 2001. Brazilian Citrus industry: economic and financial effects, 1996-2000. *Rev. Bras. Frutic.* 23:432-436.
133. **Newman, K. L., Almeida, R. P., Purcell, A. H. e Lindow, S. E.** 2003. Use of a green fluorescent strain for analysis of *Xylella fastidiosa* colonization of *Vitis vinifera*. *Appl Environ Microbiol* 69:7319-27.
134. **Newman, K. L., Almeida, R. P., Purcell, A. H. e Lindow, S. E.** 2004. Cell-cell signaling controls *Xylella fastidiosa* interactions with both insects and plants. *Proc Natl Acad Sci U S A* 101:1737-42.
135. **Newman, K. L., Chatterjee, S., Ho, K. A. e Lindow, S. E.** 2008. Virulence of plant pathogenic bacteria attenuated by degradation of fatty acid cell-to-cell signaling factors. *Mol Plant Microbe Interact* 21:326-34.
136. **Ng, W. L., Kazmierczak, K. M., Robertson, G. T., Gilmour, R. e Winkler, M. E.** 2003. Transcriptional regulation and signature patterns revealed by microarray analyses of *Streptococcus pneumoniae* R6 challenged with sublethal concentrations of translation inhibitors. *J Bacteriol* 185:359-70.
137. **Nogaroto, V., Tagliavini, S., Gianotti, A., Mikawa, A., Barros, N. T., Puzer, L., Carmona, A. K., Costa, P. e Henrique-Silva, F.** 2006. Recombinant expression and characterization of a *Xylella fastidiosa* cysteine protease differentially expressed in a nonpathogenic strain. *FEMS Microbiol Lett* 261:187-93.
138. **Nunes, L. R., Rosato, Y. B., Muto, N. H., Yanai, G. M., da Silva, V. S., Leite, D. B., Goncalves, E. R., de Souza, A. A., Coletta-Filho, H. D., Machado, M. A., Lopes, S. A. e de Oliveira, R. C.** 2003. Microarray analyses of *Xylella fastidiosa* provide evidence of coordinated transcription control of laterally transferred elements. *Genome Res* 13:570-8.
139. **Ogle, J. M., Brodersen, D. E., Clemons, W. M., Jr., Tarry, M. J., Carter, A. P. e Ramakrishnan, V.** 2001. Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science* 292:897-902.
140. **Ogle, J. M., Carter, A. P. e Ramakrishnan, V.** 2003. Insights into the decoding mechanism from recent ribosome structures. *Trends Biochem Sci* 28:259-66.
141. **Osiro, D., Colnago, L. A., Otononi, A. M., Lemos, E. G., de Souza, A. A., Coletta Filho, H. D. e Machado, M. A.** 2004. A kinetic model for *Xylella fastidiosa* adhesion,

- biofilm formation, and virulence. FEMS Microbiol Lett 236:313-8.
142. **Overhage, J., Campisano, A., Bains, M., Torfs, E. C., Rehm, B. H. e Hancock, R. E.** 2008. Human host defense peptide LL-37 prevents bacterial biofilm formation. Infect Immun 76:4176-82.
143. **Pape, T., Wintermeyer, W. e Rodnina, M. V.** 2000. Conformational switch in the decoding region of 16S rRNA during aminoacyl-tRNA selection on the ribosome. Nat Struct Biol 7:104-7.
144. **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 9:77-90.
145. **Petersen, D., Chandramouli, G. V., Geoghegan, J., Hilburn, J., Paarlberg, J., Kim, C. H., Munroe, D., Gangi, L., Han, J., Puri, R., Staudt, L., Weinstein, J., Barrett, J. C., Green, J. e Kawasaki, E. S.** 2005. Three microarray platforms: an analysis of their concordance in profiling gene expression. BMC Genomics 6:63.
146. **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.
147. **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). Int J Syst Evol Microbiol 58:2147-57.
148. **Poly, F., Read, T., Tribble, D. R., Baqar, S., Lorenzo, M. e Guerry, P.** 2007. Genome sequence of a clinical isolate of *Campylobacter jejuni* from Thailand. Infect Immun 75:3425-33.
149. **Pooler, M. R. e Hartung, J. S.** 1995. Specific PCR detection and identification of *Xylella fastidiosa* strains causing citrus variegated chlorosis. Curr Microbiol 31:377-81.
150. **Pooler, M. R., Ritchie, D. F. e Hartung, J. S.** 1996. Genetic relationships among strains of *Xanthomonas fragariae* based on random amplified polymorphic DNA PCR, repetitive extragenic palindromic PCR, and enterobacterial repetitive intergenic consensus PCR data and generation of multiplexed PCR primers useful for the identification of this phytopathogen. Applied and Environmental Microbiology 62:3121-3127.
151. **Poon, W. W., Davis, D. E., Ha, H. T., Jonassen, T., Rather, P. N. e Clarke, C. F.** 2000. Identification of *Escherichia coli* *ubiB*, a gene required for the first monooxygenase step in ubiquinone biosynthesis. J Bacteriol 182:5139-46.
152. **Purcell, A. H. e Hopkins, D. L.** 1996. Fastidious xylem-limited bacterial plant pathogens. Annual Review of Phytopathology 34:131-151.
153. **Purcell, A. H. e Saunders, S. R.** 1999. Fate of Pierce's disease strains of *Xylella fastidiosa* in common riparian plants in California. Plant Disease 83:825-830.
154. **Purcell, A. H., Saunders, S. R., Hendson, M., Grebus, M. E. e Henry, M. J.** 1999. Causal role of *Xylella fastidiosa* in oleander leaf scorch disease. Phytopathology 89:53-58.

155. **Purcino, R. P., Medina, C. L., Martins de Souza, D., Winck, F. V., Machado, E. C., Novello, J. C., Machado, M. A. e Mazzafera, P.** 2007. *Xylella fastidiosa* disturbs nitrogen metabolism and causes a stress response in sweet orange *Citrus sinensis* cv. Pera. J Exp Bot 58:2733-44.
156. **Qiu, J., Zhou, D., Han, Y., Zhang, L., Tong, Z., Song, Y., Dai, E., Li, B., Wang, J., Guo, Z., Zhai, J., Du, Z., Wang, X. e Yang, R.** 2005. Global gene expression profile of *Yersinia pestis* induced by streptomycin. FEMS Microbiol Lett 243:489-96.
157. **Quackenbush, J.** 2002. Microarray data normalization and transformation. Nat Genet 32 Suppl:496-501.
158. **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. Appl Environ Microbiol 75:5631-8.
159. **Redak, R. A., Purcell, A. H., Lopes, J. R., Blua, M. J., Mizell, R. F. e Andersen, P. C.** 2004. The biology of xylem fluid-feeding insect vectors of *Xylella fastidiosa* and their relation to disease epidemiology. Annu Rev Entomol 49:243-70.
160. **Reddy, J. D., Reddy, S. L., Hopkins, D. L. e Gabriel, D. W.** 2007. TolC is required for pathogenicity of *Xylella fastidiosa* in *Vitis vinifera* grapevines. Mol Plant Microbe Interact 20:403-10.
161. **Rodnina, M. V. e Wintermeyer, W.** 2001. Ribosome fidelity: tRNA discrimination, proofreading and induced fit. Trends Biochem Sci 26:124-30.
162. **Rodrigues, J. L., Silva-Stenico, M. E., Gomes, J. E., Lopes, J. R. 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. Appl Environ Microbiol 69:4249-55.
163. **Roper, M. C., Greve, L. C., Labavitch, J. M. e Kirkpatrick, B. C.** 2007a. Detection and visualization of an exopolysaccharide produced by *Xylella fastidiosa* in vitro and in planta. Appl Environ Microbiol 73:7252-8.
164. **Roper, M. C., Greve, L. C., Warren, J. G., Labavitch, J. M. e Kirkpatrick, B. C.** 2007b. *Xylella fastidiosa* requires polygalacturonase for colonization and pathogenicity in *Vitis vinifera* grapevines. Molecular Plant-Microbe Interactions 20:411-419.
165. **Rosato, Y. B., Neto, J. R., Miranda, V. S., Carlos, E. F. e Manfio, G. P.** 1998. Diversity of a *Xylella fastidiosa* population isolated from *Citrus sinensis* affected by Citrus Variegated Chlorosis in Brazil. Systematic and Applied Microbiology 21:593-598.
166. **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.
167. **Rozen, S. e Skaletsky, H.** 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132:365-86.
168. **Rudrappa, T., Biedrzycki, M. L. e Bais, H. P.** 2008. Causes and consequences of plant-associated biofilms. Fems Microbiology Ecology 64:153-166.
169. **Ryan, R. P., Monchy, S., Cardinale, M., Taghavi, S., Crossman, L., Avison, M. B., Berg, G., van der Lelie, D. e Dow, J. M.** 2009. The versatility and adaptation of

- bacteria from the genus *Stenotrophomonas*. *Nat Rev Microbiol* 7:514-25.
170. **Scally, M., Schuenzel, E. L., Stouthamer, R. e Nunney, L.** 2005. Multilocus sequence type system for the plant pathogen *Xylella fastidiosa* and relative contributions of recombination and point mutation to clonal diversity. *Applied and Environmental Microbiology* 71:8491-8499.
171. **Scarpari, L. M., Lambais, M. R., Silva, D. S., Carraro, D. M. e Carrer, H.** 2003. Expression of putative pathogenicity-related genes in *Xylella fastidiosa* grown at low and high cell density conditions in vitro. *FEMS Microbiol Lett* 222:83-92.
172. **Schaad, N. W., Postnikova, E., Lacy, G., Fatmi, M. e Chang, C. J.** 2004. *Xylella fastidiosa* subspecies: *X. fastidiosa* subsp. *piercei*, subsp. nov., *X. fastidiosa* subsp. *multiplex* subsp. nov., and *X. fastidiosa* subsp. *pauca* subsp. nov. *Syst Appl Microbiol* 27:290-300.
173. **Schmid, A. K., Reiss, D. J., Pan, M., Koide, T. e Baliga, N. S.** 2009. A single transcription factor regulates evolutionarily diverse but functionally linked metabolic pathways in response to nutrient availability. *Mol Syst Biol* 5:282.
174. **Schoolnik, G. K.** 2002a. Functional and comparative genomics of pathogenic bacteria. *Curr Opin Microbiol* 5:20-6.
175. **Schoolnik, G. K.** 2002b. Microarray analysis of bacterial pathogenicity. *Adv Microb Physiol* 46:1-45.
176. **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*. *Appl Environ Microbiol* 71:3832-9.
177. **Setubal, J. C., Moreira, L. M. e da Silva, A. C.** 2005. Bacterial phytopathogens and genome science. *Curr Opin Microbiol* 8:595-600.
178. **Shaw, K. J., Miller, N., Liu, X., Lerner, D., Wan, J., Bittner, A. e Morrow, B. J.** 2003. Comparison of the changes in global gene expression of *Escherichia coli* induced by four bactericidal agents. *J Mol Microbiol Biotechnol* 5:105-22.
179. **Shi, X., Bi, J., Morse, J. G., Toscano, N. C. e Cooksey, D. A.** 2009a. Differential expression of genes of *Xylella fastidiosa* in xylem fluid of citrus and grapevine. *FEMS Microbiol Lett*.
180. **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*. *Appl Environ Microbiol* 73:6748-56.
181. **Shi, X. Y., Dumenyo, C. K., Hernandez-Martinez, R., Azad, H. e Cooksey, D. A.** 2009b. Characterization of regulatory pathways in *Xylella fastidiosa*: genes and phenotypes controlled by *gacA*. *Appl Environ Microbiol* 75:2275-83.
182. **Shoji, S., Walker, S. E. e Fredrick, K.** 2006. Reverse translocation of tRNA in the ribosome. *Mol Cell* 24:931-42.
183. **Silva-Stenico, M. E., Pacheco, F. T., Rodrigues, J. L., Carrilho, E. e Tsai, S. M.** 2005. Growth and siderophore production of *Xylella fastidiosa* under iron-limited conditions. *Microbiol Res* 160:429-36.
184. **Silva-Stenico, M. E., Pacheco, F. T. H., Pereira, E. R., Rodrigues, J. L. M., Souza, A. N., Etchegaray, A., Gomes, J. E. e Tsai, S. M.** 2009. Nutritional deficiency in citrus with symptoms of citrus variegated chlorosis disease. *Brazilian Journal of Biology* 69:861-866.

185. **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* 406:151-7.
186. **Souza, L. C., Wulff, N. A., Gaurivaud, P., Mariano, A. G., Virgilio, A. C., Azevedo, J. L. e Monteiro, P. B.** 2006. Disruption of *Xylella fastidiosa* CVC gumB and gumF genes affects biofilm formation without a detectable influence on exopolysaccharide production. *FEMS Microbiol Lett* 257:236-42.
187. **Stenger, D. e Chen, J.** 2008. *Xylella fastidiosa* isolates from mulberry harbor a 25 kilobase pair plasmid with extensive sequence identity to a plasmid from *Verminephrobacter eiseniae*. *Phytopathology* 98:S150-S150.
188. **Stiens, M., Becker, A., Bekel, T., Godde, V., Goesmann, A., Niehaus, K., Schneiker-Bekel, S., Selbitschka, W., Weidner, S., Schluter, A. e Puhler, A.** 2008. Comparative genomic hybridisation and ultrafast pyrosequencing revealed remarkable differences between the *Sinorhizobium meliloti* genomes of the model strain Rm1021 and the field isolate SM11. *J Biotechnol* 136:31-7.
189. **Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. e Higgins, D. G.** 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25:4876-82.
190. **Thompson, J. D., Higgins, D. G. e Gibson, T. J.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673-80.
191. **Travensolo, R. F., Carareto-Alves, L. M., Costa, M. V. C. G., Lopes, T. J. S., Carrilho, E. e Lemos, E. G.** 2009. *Xylella fastidiosa* gene expression analysis by DNA microarrays. *Genetics and Molecular Biology* 32:340-353.
192. **Twellmeyer, J., Wende, A., Wolfertz, J., Pfeiffer, F., Panhuysen, M., Zaigler, A., Soppa, J., Welzl, G. e Oesterhelt, D.** 2007. Microarray analysis in the archaeon *Halobacterium salinarum* strain R1. *PLoS One* 2:e1064.
193. **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*. *J Bacteriol* 185:1018-26.
194. **Van Sluys, M. A., Monteiro-Vitorello, C. B., Camargo, L. E., Menck, C. F., Da Silva, A. C., Ferro, J. A., Oliveira, M. C., Setubal, J. C., Kitajima, J. P. e Simpson, A. J.** 2002. Comparative genomic analysis of plant-associated bacteria. *Annu Rev Phytopathol* 40:169-89.
195. **VanBogelen, R. A. e Neidhardt, F. C.** 1990. Ribosomes as sensors of heat and cold shock in *Escherichia coli*. *Proc Natl Acad Sci U S A* 87:5589-93.
196. **Venkatasubbarao, S.** 2004. Microarrays--status and prospects. *Trends Biotechnol* 22:630-7.
197. **Vidaver, A. K.** 2002. Uses of antimicrobials in plant agriculture. *Clinical Infectious Diseases* 34:S107-S110.
198. **Vollmer, W., Joris, B., Charlier, P. e Foster, S.** 2008. Bacterial peptidoglycan (murein) hydrolases. *FEMS Microbiol Rev* 32:259-86.
199. **Wang, L. H., He, Y., Gao, Y., Wu, J. E., Dong, Y. H., He, C., Wang, S. X., Weng, L. X., Xu, J. L., Tay, L., Fang, R. X. e Zhang, L. H.** 2004. A bacterial cell-cell communication signal with cross-kingdom structural analogues. *Mol Microbiol* 51:903-

- 12.
200. **Warren, J. G. e Kirkpatrick, B. C.** 2009. Inhibition of *Xylella fastidiosa* polygalacturonase to produce Pierce's disease resistant grapevines. *Phytopathology* 99:S138-S138.
201. **Wickert, E., Machado, M. A. e Lemos, E. G.** 2007. Evaluation of the Genetic Diversity of *Xylella fastidiosa* Strains from Citrus and Coffee Hosts by Single-Nucleotide Polymorphism Markers. *Phytopathology* 97:1543-9.
202. **Woese, C. R.** 1987. Bacterial evolution. *Microbiol Rev* 51:221-71.
203. **Wolf, D., Gray, C. P. e de Saizieu, A.** 2000. Visualising gene expression in its metabolic context. *Brief Bioinform* 1:297-304.
204. **Zaini, P. A.** 2007. O stimulon de ferro em *Xylella fastidiosa*. Tese de Doutorado. Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, São Paulo. 101.
205. **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. *J Bacteriol* 190:2368-78.
206. **Zengel, J. M. e Lindahl, L.** 1994. Diverse mechanisms for regulating ribosomal protein synthesis in *Escherichia coli*. *Prog Nucleic Acid Res Mol Biol* 47:331-70.
207. **Zhang, Z., Pendse, N. D., Phillips, K. N., Cotner, J. B. e Khodursky, A.** 2008. Gene expression patterns of sulfur starvation in *Synechocystis* sp. PCC 6803. *BMC Genomics* 9:344.

Anexos

Tab elas Suplementares

Tabela S1. Comparação do perfil de expressão das CDS moduladas negativamente pela estreptomicina aos seus respectivos perfis quando *X. fastidiosa* é submetida a outras condições experimentais.

Tabela S2. Comparação do perfil de expressão das CDS anotadas no genoma de *X.fastidiosa* 9a5c em diversas condições experimentais. (disponível no CD-ROM suplementar a esta Tese).

Tabela S3. Comparação do perfil de expressão das CDS selecionadas em diversas condições experimentais.

Tabela S4. Produto predito das CDS selecionadas em diversas condições experimentais e listadas na Tabela S3.

Tabela S5. Resultado do mapeamento dos *reads* do sequenciamento do genoma da cepa J1a12 no genoma referência (cepa 9a5c). (disponível no CD-ROM suplementar a esta Tese).

Figura Suplementar

Figura S1. Alinhamento das sequências de nucleotídeos do *contig* 374 e do plasmídeo pVEIS01. As sequências foram alinhadas utilizando-se o programa CLUSTAL. (*) indica nucleotídeos idênticos. (Figura completa está disponível no CD-ROM suplementar a esta Tese).

Material Suplementar

Pirosequenciamento do DNA genômico de *X. fastidiosa*

Tabela S1. Comparação do perfil de expressão das CDS moduladas negativamente pela estreptomicina aos seus respectivos perfis quando *X. fastidiosa* é submetida a outras condições experimentais.

CDS	STR	GOM	HIC	LIC	HS	SS	OS	Glu1	Glu250	FP	BF	BF2	PW/XDM	J1a12	rpoN	BCYE/ XDM2	PD gacA	PD X1/X2
XF0012	-0,9					-1,5	-1,4											
XF0068	-2,0			-1,6									-1,1					
XF0073	-0,7																	
XF0083	-1,6		-1,7	-0,7	-1,5		-1,0	-0,7							1,1			1,4
XF0104	-2,4																	
XF0112	-2,0					1,2	2,1											
XF0121	-1,5			0,8														
XF0145	-1,8																	
XF0246	-1,9																	
XF0274	-1,3	1,5	-3,8		-0,9	-1,5	-2,4									-1,8		
XF0286	-3,4		-3,0	-1,4														
XF0298	-1,7												1,8					
XF0307	-1,0																	
XF0320	-0,8			1,2	-0,8		-1,2									-1,4		
XF0321	-1,4											-1,4	1,1				-2,1	
XF0323	-1,2	1,7				3,1	4,5											
XF0328	-2,3				0,7		1,2											
XF0339	-0,9	2,0												1,6				
XF0340	-1,3		0,9				-1,4											
XF0345	-1,5		1,2															
XF0357	-1,0									1,7						-0,6	2,7	
XF0359	-1,3																	
XF0409	-1,7																	
XF0411	-0,9																	
XF0413	-2,4																	
XF0474	-3,1																	

XF0475	-3,1																	
XF0480	-1,6	5,3																
XF0483	-1,7				0,5													
XF0485	-1,1																	
XF0488	-1,1					0,9												
XF0497	-1,7				-1,0					-0,7					0,7			
XF0584	-1,3						-1,0											
XF0632	-0,8									0,7		1,0						
XF0766	-1,3					1,7	2,6								-0,5			
XF0774	-1,1																	
XF0807	-0,9		1,9		1,8													
XF0822	-1,1														0,8			
XF0838	-0,6																	
XF0852	-2,7																	
XF0870	-2,1					1,1												
XF0881	-2,1	1,6			1,9										-1,0			
XF0885	-1,2	2,7																
XF0890	-2,5	1,4			-0,5							1,7						
XF0902	-2,3	2,0																
XF0945	-1,8																	
XF1028	-1,0		2,5	2,9														
XF1037	-2,0					-1,1	-1,6											
XF1061	-0,7																	
XF1090	-1,0																	
XF1116	-0,8														-2,2			
XF1122	-1,6																	
XF1129	-0,7		-1,5	-1,1														
XF1139	-1,7																	
XF1140	-3,3				0,9		1,1											
XF1143	-0,9		-1,3	-1,1														
XF1188	-1,0									2,1								

XF1226	-0,8			0,6			-1,3											
XF1243	-1,8						1,5											
XF1244	-0,6									1,8								
XF1246	-1,2						1,4											
XF1275	-1,0		-2,9	-1,2	-0,9													
XF1374	-1,8			1,1	2,1					0,9								
XF1408	-1,5		-2,5	-2,6	-1,0													
XF1506	-2,0																	
XF1516	-1,4		-2,0		1,0		1,0			3,3		-1,7					-3,2	1,2
XF1589	-1,9		-1,8			1,1	2,8											
XF1604	-0,8									1,4								
XF1620	-1,0														0,7			
XF1629	-0,4									1,8								
XF1661	-3,8					0,9	2,2		1,5									
XF1737	-2,7												2,1					
XF1738	-0,7												2,5					
XF1760	-1,0												1,6					
XF1806	-1,4												-1,0					
XF1812	-1,7	6,9															-3,0	
XF1835	-2,2				-0,7													
XF1855	-1,8						-1,3										-2,0	
XF1884	-0,7			-1,6			-2,2											
XF1887	-2,3		-1,4	-0,7			-1,1											
XF1911	-0,9																	
XF2012	-0,6																	
XF2017	-0,7																	
XF2022	-2,5	1,8		-1,8														
XF2049	-3,4																	
XF2050	-1,2				-0,4							-1,2						
XF2144	-0,7		-2,5	-1,6			-1,0											
XF2245	-4,6						-1,6											

XF2318	-0,9				0,7													
XF2366	-0,8	2,8																
XF2370	-1,5																	
XF2430	-0,6	2,8		1,3														
XF2434	-0,7			1,0														
XF2448	-1,7																	
XF2463	-1,1												1,5					
XF2472	-0,9												1,8					
XF2477	-2,4												1,5					
XF2514	-0,9		-1,9		2,6	2,2	3,4											
XF2547	-1,2		-2,4	-1,4	-1,7													
XF2548	-1,7		-1,9		-1,3									-0,6				
XF2552	-1,7																	
XF2687	-1,1															-1,7		
XF2688	-1,1									-0,6						-2,3		
XF2703	-0,7																	
XF2753	-1,6	3,0				1,3												
XFa0006	-0,5		-0,8															
XFa0009	-0,9				-1,0													
XFa0028	-0,9				1,2	1,3	2,8				1,9					-1,6		
XFa0033	-1,7						1,9		0,7									

As siglas no topo de cada coluna indicam as diferentes condições ou tratamentos em que a expressão gênica de *X. fastidiosa* foi avaliada. Os dados (valor M) da coluna STR foram obtidos como descrito na Tabela 9. Os demais dados (valor M) foram extraídos da literatura como descrito no item 4.2. STR (estreptomicina), GOM (gomesina), HIC (excesso de ferro), LIC (ausência de ferro), HS (choque térmico), SS (sacarose 300mM), OS (NaCl 250mM), Glu1 (Glicose 1mM), Glu250 (Glicose 250mM), BF (biofilme), BF2 (biofilme em lenho), PW/XDM (meio PW versus XDM), J1a12 (cepa J1a12 versus 9a5c), rpoN (mutante de *rpoN*); BCYE/XDM2 (meio BCYE versus XDM2), PD-gacA (mutante de *gacA*) e PD X1/X2 (cepa PD xilema de videira versus de citros).

Tabela S3. Comparação do perfil de expressão das CDS selecionadas em diversas condições experimentais.

CDS	STR	GOM	HIC	LIC	HS	SS	OS	Glu1	Glu250	FP	BF	BF2	PW/ XDM	J1a12	rpoN	rpoE	BCYE/ XDM2	algU_ PD	gacA_ PD	PD X1/X2
XF0089				0,9														-1,6		1,4
XF0155			3,0	1,7		1,8								1,8						
XF0172			2,5	1,2			2,4							-1,6			-2,5			
XF0261			1,0	0,8					1,2											
XF0262			0,7	1,3	-1,1				1,3										1,9	
XF0263			0,4		-1,4	0,9				2,7		-2,5	-1,1							
XF0264				1,5	-1,2		1,3		1,3									-1,3	-1,0	
XF0304					0,8															1,3
XF0309						-1,2					0,8									
XF0449			1,7								0,7			1,5						
XF0479												-2,2								1,1
XF0649			1,3		1,5															
XF0667			0,8	1,2	1,0	2,3			0,9											
XF0785			1,0	0,9	0,9						0,9						-2,4			
XF0832		2,5				-1,3	1,2				0,7		2,0							
XF0833		1,5					1,2						1,4							
XF0989						1,1			7,0											
XF1024			-1,4	-1,4		-2,5			-1,1					2,1						1,3
XF1053			1,5				2,2													
XF1056			2,1	1,3	2,0								-1,4	-2,4						
XF1187			-1,4		0,6			1,1												0,7
XF1212			1,4													1,7				
XF1216		2,9	1,8		-1,6									1,4						
XF1220			1,6			1,4	2,3										-0,7			
XF1255				2,2						3,0								-1,2	1,7	
XF1306			2,5								0,8									

XF1318			1,1	1,3	0,9														
XF1350			1,3							0,9									
XF1416								7,0		2,0									
XF1482				1,0						1,2									
XF1493			1,9			1,5	1,6												
XF1516	-1,4		-2,0		1,0		1,0		3,3		-1,7						-3,2	1,2	
XF1529			-2,2	-2,2					1,9		-2,1						-1,7	0,8	
XF1614							1,2			0,8									
XF1632			-2,8	-1,3	-0,9											-0,7		0,8	
XF1633			-1,8									-1,1						0,7	
XF1721		3,9			1,6	-1,0						2,2							
XF1739		1,5										1,7							
XF1755			2,1	1,2			3,3					2,4							
XF1756				0,8		1,3	3,3			0,8		2,8							
XF1937		1,3								1,2						0,9			
XF1940							1,4		0,8									0,7	
XF1954			0,8															0,8	
XF1955			1,4															0,9	
XF2071			1,3	0,6	1,0	1,1													
XF2177		1,7	1,3	1,1															
XF2195			2,0	1,7	0,5	2,0													
XF2196		1,3		0,8	2,2														
XF2239				-1,5														0,8	
XF2257					0,8	2,2	3,8			2,0									
XF2336					2,0								1,3	1,1		-0,7			
XF2338					3,0					1,1									
XF2364		2,7																	
XF2367		2,9																	
XF2369		3,2										-1,1					1,7		
XF2377			0,6	0,7	1,7			1,0	1,0				-1,3						
XF2519						2,6	2,3			1,0									
XF2535			1,2	1,2			1,3												

XF2594				0,6	3,9											4,0				
XF2677							1,9						-1,2						1,5	
XF2715				1,4	1,4	1,3														
XF2781			1,8	1,6		0,8	1,4													

As siglas no topo de cada coluna indicam as diferentes condições ou tratamentos em que a expressão gênica de *X. fastidiosa* foi avaliada. O nome do produto predito está listado na Tabela S4. Os dados (valor M) da coluna STR foram obtidos como descrito na Tabela 9. Os demais dados (valor M) foram extraídos da literatura como descrito no item 4.2. e foram extraídos da Tabela S2. STR (estreptomicina), GOM (gomesina), HIC (excesso de ferro), LIC (ausência de ferro), HS (choque térmico), SS (sacarose 300mM), OS (NaCl 250mM), Glu1 (Glicose 1mM), Glu250 (Glicose 250mM), BF (biofilme), BF2 (biofilme em lenho), PW/XDM (meio PW versus XDM), J1a12 (cepa J1a12 versus 9a5c), rpoN (mutante de *rpoN*); BCYE/XDM2 (meio BCYE versus XDM2), PD-gacA (mutante de *gacA*) e PD X1/X2 (cepa PD xilema de videira versus de citros).

Tabela S4. Produto predito das CDS selecionadas em diversas condições experimentais e listadas na Tabela S3.

CDS	Produto predito
XF0089	RNA-binding protein Hfq
XF0155	hypothetical protein
XF0172	hypothetical protein
XF0261	colicin V immunity protein
XF0262	colicin V precursor
XF0263	colicin V precursor
XF0264	colicin V precursor
XF0304	preprotein translocase subunit SecG
XF0309	NADH dehydrogenase subunit E
XF0449	hypothetical protein
XF0479	type IV pilin
XF0649	peptidase family S54 protein (rhomboid like protein)
XF0667	hypothetical protein
XF0785	sulfur deprivation response regulator
XF0832	siroheme synthase
XF0833	transcriptional regulator CysB-like protein
XF0989	conserved hypothetical protein
XF1024	outer membrane protein H8 precursor
XF1053	outer membrane protein
XF1056	conserved hypothetical protein
XF1187	ATP-dependent Clp protease proteolytic subunit

XF1212	peptidyl-prolyl cis-trans isomerase
XF1216	colicin V secretion protein
XF1220	colicin V secretion ABC transporter ATP-binding protein
XF1255	hypothetical protein
XF1306	hypothetical protein
XF1318	Xanthomonadales conserved hypothetical
XF1350	RNA polymerase sigma factor RpoD
XF1416	hypothetical protein
XF1482	hypothetical protein
XF1493	virulence regulator
XF1516	surface-exposed outer membrane protein
XF1529	surface protein
XF1614	penicillin binding protein
XF1632	twitching motility protein
XF1633	twitching motility protein
XF1721	transcriptional regulator (LysR family)
XF1739	outer membrane protein
XF1755	conserved hypothetical protein
XF1756	hypothetical protein
XF1937	proton glutamate symport protein
XF1940	methionine sulfoxide reductase A

XF1954	pilus biogenesis protein
XF1955	pilus protein
XF2071	CopG/Arc/MetJ family transcriptional regulator
XF2177	hypothetical protein
XF2195	conserved hypothetical protein
XF2196	hemagglutinin-like secreted protein
XF2239	RNA polymerase sigma factor RpoE
XF2257	predicted membrane protein DUF204
XF2336	two-component system, regulatory protein
XF2338	prephenate dehydrogenase
XF2364	GumH protein
XF2367	GumD protein
XF2369	GumC protein
XF2377	hypothetical protein
XF2519	hypothetical protein
XF2535	two-component system, sensor protein
XF2594	Peptidase M48 family protein
XF2677	Multicopper oxidase-related protein
XF2715	transcriptional regulator protein, TetR
XF2781	ribonuclease P

Figura S1 (parcial). Alinhamento das sequências de nucleotídeos do *contig* 374 e do plasmídeo pVEIS01. As sequências foram alinhadas utilizando-se o programa CLUSTAL (*) indica nucleotídeos idênticos. O alinhamento completo está disponível no CD-ROM suplementar a esta Tese.

```

CLUSTAL 2.0.12 multiple sequence alignment
pVEIS01          -----AGCGATTGCTCGTGTG 2708
374_J1a12      TATTAGTTTTCTAAAAAGTCGCGAAATATTAAGTAAGTATTATAACGATTGCTCGTGTG 11689
                  * * * * *

pVEIS01          TATCGCGCTTTTGCACGCGCTCAGTGGTCTTGTCTTAACTCTCGGCTCCGTACCTGAT 2768
374_J1a12      TATCACGCTGTTGCACGCTCGTTCAGTGGTCTTGTCTGAATCTTCGGCTCCGTTCCTGAT 11749
                  **** * * * * *

pVEIS01          TCAGGCGCTTCAAGAACTGCGGGTCAACCGCTTGTCTTGTGCAAGCGGGCGATGGCCTCAT 2828
374_J1a12      TCAGGCGTTTCAAAAAGTCGCGGTCAGCGGCTGTTTGTCAAACGGGCGATGGCCTCAT 11809
                  * * * * *

pVEIS01          CCTGCACCGGCTTGGGCTGCAACTTGTCTCTTCGGCCACGATGCCGCGCCAGTACGCCA 2888
374_J1a12      CCTGCACAGCCTTGGGCTGCAACTTGTCTCTTCGGCCACAATGCCCGCCAGTAGGCCA 11869
                  * * * * *

pVEIS01          GCTTGGTCAAGTCTTCCACGACTTGTGCGCCATGCCCGGCTGCTTTTGCAGCGCCTGCA 2948
374_J1a12      GTTGGTGCAGATCCTCTACAGACTTGTGAGCCATGTTGGGTGTTTGCAGCGCCTGCA 11929
                  * * * * *

pVEIS01          TGTGGTGGCGATGAACTTGTGCGCCTCAACGTGCATGGCCGCTTGGTCTTGTTCAGGT 3008
374_J1a12      TATTGGTAGCGATGAACTTGTGCGCCTCAACGTGCATGGCCGCTGATTTTGTTCATGT 11989
                  * * * * *

pVEIS01          CATGGCGGGGTGCGGCAGGCGTGGTTTCTTCGCGCACGGTTTCGCGCTCTGCCTTGCCTG 3068
374_J1a12      CATGGCGTGGTGCCGAGCCGTACTC-----GCCTGACGTTCTGTCTTGCCTG 12037
                  * * * * *

pVEIS01          CTGCGAGCGGGGTGATGGCGTTGGTGGAGCGCTCTTGTGCGAGAGTCTCCAGCGCGGCCA 3128
374_J1a12      CTGCGAGTGGGTGATGGAGTTCTGGAACGCTCTTGTGCGAGAGTTTCCAGTGGCGGCCA 12097
                  * * * * *

pVEIS01          AATCCGCTTGCTTGGGCGTATAGCCCTGCGTCTTGATACCTTGCAGCTCAGCTTGCAACC 3188
374_J1a12      AATCGGTTTGCTTGGTGTGTATCCCTGAGTCTTGATGCCTTGTGACTCAGCTTGCAACC 12157
                  * * * * *

pVEIS01          AGGCTTCGCGCCTGAACTCCTGCGAGCCGACAGCTTGAGACTGTCCCACTGCTTCGCTT 3248
374_J1a12      AGGCTTCACGTCGGAACCTTGTGAGCCTGACAACTTGAGACTATCCCACTGTTTCGCTT 12217
                  * * * * *

pVEIS01          TGGCGTAGGCCACCATATCGGCCACGGCTTCTTGTGGTGGTGGACGTGACCAGCTTTT 3308
374_J1a12      TGGCGTAGGCCACCATATCGGCTACAGCGTCTTTATTGGTGGTGGACGTGGCCAGCTTTT 12277
                  * * * * *

pVEIS01          CGCCCTTGCTCTCGAACATCACGCGGTTGAGTCTTGGCGTAAACCTTGCCGTCCATCT 3368
374_J1a12      CACCCTTGCTCTCAAACATCACGCGGTTTGAATCCTTGGCATAAAATTTGCCGTCCACCT 12337
                  * * * * *

pVEIS01          CGGTGTATGCCTGCGTGATGCGCGCGGCACGATGAAATCCCGCGCTCGAGCTCCTTGC 3428
374_J1a12      CGGTGTAAGTCTGAGTGATACGAGGACGATGAAATCGCCGCGTTCAAGCTCCTTGC 12397
                  * * * * *

pVEIS01          CTGAGCGCCGC---ACGTCCACCGAGTTCAAGCCCTGGCCAGCTTGCTCCTGTTACGCG 3485
374_J1a12      CTGTGCGTTGTTGTACGTCTACAGAATTCAAGCCTTGGCCAGCCTGTTCTGTTACGCG 12457
                  *** * * * *

pVEIS01          CCTGCGGCGGCGCAAATCGTCTTCGAGCTGGTCTTGCAAGCCCTGTTACGCGCGCGG 3545
374_J1a12      CTTTGCAGCGGCGTAGTTTCGTCTTCAAGCTGCTCTTTCAATGCTTGTGCGACAGCCGCTG 12517
                  * * * * *

pVEIS01          CTGCTCCCGGTGCTAATCACCGGCTTGGCTTCTTACCTTTGGCGGCTGCGGCATGGA 3605
374_J1a12      TTTGTTCTTCGGTTGTT----- 12534
                  *** * * * *

pVEIS01          AGGCCGTCCGCGAGGTAAACGAATCGCCTTACCCCGTTTGCCAAGCGCCACCAGGTCGG 3665
374_J1a12      -----CGAGTCACT----- 12543
                  * * * * *

pVEIS01          CATGGTCGGCCACGTCGCGGCGAGCTCGCGCCTTGGCTTGTTCCTCAGACCAGGCGCGGG 3725
374_J1a12      -----

```

```

pVEIS01      CCGTAGCCTCGCGCTCTTGCCGCATCGAATCGGCCTCTTGCTCTTTGCGCTCTTCCTTGG 3785
374_J1a12    -----ATGGACTCGGATTCTTGCGCTTTGCGATTCTTCTTTGG 12580
              ** ** *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

pVEIS01      CCTTCGCGGCCAGCAGCTCGCTATGCGTGTCTTCGGGTGCCACGCCAGGCTCGCCGCGCA 3845
374_J1a12    CCTTAGCCGTACGAGCTCGTTATACGTGTCTTCCGGCGACACAGTAGGCTCATCACGCA 12640
              **** *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

pVEIS01      GGTGTGTCGCGCAACGCCTTGTCTTGCCATCTTGCAACAGCTCGCTTCGGGTGTCTCGG 3905
374_J1a12    AGTTGTACGCAAGACTTTGCTTTGAACATCTTGTAAGCTCGCTTCGGGTGTCTCGG 12700
              *****  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

pVEIS01      GTGTACACACAGGTTACCCGCGCAGGCGGTTTCTCATGTCGGCATCGGATGCTTCGGGGG 3965
374_J1a12    GGGTCACATCAGGTTACCCGCGCAGAGGTTTTCGAGGTCGGTATTAGAGTCTTTGAGG 12760
              *  *****  *****  *  *  *  *  *  *  *  *  *  *  *  *  *

pVEIS01      TAGTCG-----CCATGTTGCAGTCTCCTTAAATTCCAAATTCCTTGAGGATGGCCAAGC 4019
374_J1a12    TGGTCGTCGTTGCCATGTGCCATCTCCTTAAATTCCAAATTCCTTGAGGATGGCCAAGC 12820
              *  ****  *****  *****  *****  *****  *****

pVEIS01      CACGGTCAACGGCCATGACCGTTTCAACCGAGGCTTTGGGTTTCAAGTGGACAGTCGCAGCC 4079
374_J1a12    CACGGTCAACGGCCATGACCGTTTCAACCGATGCTTGGGTTTCAAGTGGACAGGCGCAGCC 12880
              *****  *****  *  *  *****  *****  *****

pVEIS01      ACTCAGTCCGGTACACGCTGGCCAGCTCGGGCGCGTTGCGGAATGCCACCGCCGTTTCGG 4139
374_J1a12    ATTCCGTCCTATACAAGCTGGCCAGCTCAGGGGCAATGCGGAATGCTACCGCTGTTTCGG 12940
              *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

pVEIS01      CGTTGCGCAGATTTCGCGGAAGACGTGTAGTTGAAGCTCCCGAGTGCACATGCAGGCCGT 4199
374_J1a12    CGTTGCGTAGGTTTGCCGAAGACGTGTAGTTGAAGCTCCCGAGTGCACATGCAGGCCAT 13000
              *****  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

pVEIS01      CAGCCACCATGAACTTGTGGTGCATCGCCGCGTAGTTCTCGCACCGGAAAACCGGGATGC 4259
374_J1a12    CAGCCACCATGAACTTGTGGTGCATCGCCGCGTAGTTCTCGCACCTAAAACAGGGATAC 13060
              *****  *****  *****  *****  *****

pVEIS01      CTTCGACGACAGGTAATCGACGCGCTTATAGCCGCCCTGATCGTCGGTGTTTTGCGCGT 4319
374_J1a12    CTTCGACGACAGTAATCGACGCGCTTATAGCCGCCCTGATCGTCGGTGTTTTGCGCGT 13120
              *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

pVEIS01      GATCGACCAACCGCGAACATCAATGCGCGTGCCTTGGCCTCAACCAACGCAAAGGCAA 4379
374_J1a12    GATCAACCAACCGCGAACATCAATGCGCGTGCCTTGGCCTCAGTCAAGCAAGGCAA 13180
              ****  *****  *****  *****  *****

pVEIS01      TGTCTTGTCTGGTGAACGAGTACGCAGCCACCAGAAGTTCCGTCTTTGCGCCGAGCACAA 4439
374_J1a12    TGTCTTGTCTGGTGAACGAGTACGCAGCCACCAGTTCAGTCTTGGCACCCAGCACGA 13240
              *****  *****  *****  *****  *****  *  *  *  *  *

pVEIS01      ACGACATGACCACCTCAAGCGCTTGCCTTCGGTGAGAAGCCAACTTCAAACGTCGCCC 4499
374_J1a12    ACGACATGACCACCTCAAGGCGTTACGCTTGGGTGAGAACCCTCAAACTCAAACGTCGCCC 13300
              *****  *****  *****  *****  *****  *****

pVEIS01      CGGTGGGCAACGCTTGTCTTCAACCGCACCCGCGAAATGCAGGACAGGGGAGGGCGGTG 4559
374_J1a12    CAGTGGGTAAATGCTTGTCTTCAATCGCGCTGCAAAGTCCAGATAGGGGAGGGGGGAG 13360
              *  *****  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

pVEIS01      CAGGCTCAATCGTGTGAGCAGCTCGCTTGGTCAAGCACCGCAGGTTGCGGGAGGTTTG 4619
374_J1a12    CTGGTTCAATGTGTGACAGATCGCTTGGTCAAGAACATCAGGTTGTTGGAAGGTTTG 13420
              *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

pVEIS01      CCCGAGCCTGAAAAATCGGCTCTTGATACCAAAAGATTTGTCCACAGCGCCGGGGGCA 4679
374_J1a12    CCCTAGCCTGAAAAATCGGTTCTTGATACCAAAATATTTGTCTACAGCGCCGGAGCCA 13480
              ***  *****  *****  *****  *****  *****  *  *  *

pVEIS01      TGTCTCCATGTGATGATCTCTTGGCCATGCCTATGGCCTTCACTTCCTGGGGCAGCA 4739
374_J1a12    TGTCTCCATGTGATGATCTCTTGAATGGCCGATCGCTTCACTTCCTGGGGGAAGCA 13540
              *****  *****  *  *  *  *  *  *  *  *  *  *  *  *  *

pVEIS01      TCACAGCCCGCGCTGATCGCTTTCGACTCGCTGCGGCCGACTTCGCGCCGAGCTGGC 4799
374_J1a12    ATACCGCACGGCGTTGATCGCTTTCGATTCACTGCTCCAGACTTCAGGCCGAGTTGAC 13600
              *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

pVEIS01      GAGACTTCGACAAGCCCTTTTCGGTCAATGATCCGATCATCTCGCTGTAGTCTTCGCGT 4859
374_J1a12    GAGACTTAGATAAACCTTCTCGGTCAATGATCCGATCATCTCGCTGTAAATCCTTTGCAT 13660
              *****  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *  *

```

pVEIS01
374_J1a12 CGTCATCGTCGCGGGGGGCATAGACAACCTGCAACGCCATGTTGGTCAACATGGTTTGAG 4919
CATCGTCGTCCCGAGGGGCATAGGCAACCTGCAAGGCCATGTTGTCAACATGGTTTGAG 13720
* * * * *

pVEIS01
374_J1a12 CGCCGGTTTCGTCGTACAGCTTGCGGTCTTTGACCTGGCTCTTGCTCTGAAAAACCAGCA 4979
CGCCGGTTTCGTCGTATAGCTTGCGGTCTTTAATCTGGCTTTTGCTTTGGAAAAATCAGCA 13780
* * * * *

pVEIS01
374_J1a12 GCAAGCGCATGTTGTACCCGGCCATGTACGCCGATGCCTTCTCAATGATGGCGATGCGCC 5039
GCAAGCGCATGTTGTACCCAGCCATGTACGCCGATGCCTTTTCAATGATGGCGATGCGCC 13840
* * * * *

pVEIS01
374_J1a12 CCATCGCCGGGAACCTCATCGAGCAGCAACAGGCATTGGTATTTCAGGCTGGGGTCATCCT 5099
CCATGGCAGGGAACTCATCAAGCAGCAACAGGCATTGGTACCTCAAGCTAGGGTCATCCT 13900
* * * * *

pVEIS01
374_J1a12 GCGGCAGCTTGTCCATGTTTCAGCACGAGAAGCTGTTGGAAAAACAGGTTCAAGAACTTGC 5159
GCGGCAGCTTGTCCATGTTTCAGTACAAGAAGTTGTTGAAAAAAAGATTCAAAAACCTTAC 13960
* * * * *

pVEIS01
374_J1a12 CGAACTTGTCAATGTTGGCGGGCTTGACCACCACGTAGATGCTCATGCGCTTTTTCGCGCA 5219
TGAACCTTGTCAATATTGGCAGGCTTCACCACCACATAGATGCTCATGCGCTTTTTCGCGCA 14020
* * * * *

pVEIS01
374_J1a12 CGTCGCGCATATCGAAGCTATTGAGGCCGTGGCCTTTGCAGTGATTGGGTCACTGAAAA 5279
CGTCGCGCATATCGAAGCTGTTAGAGGCCGTGCTTTTGCAGTGATCGGGTCACTGAAGA 14080
* * * * *

pVEIS01
374_J1a12 TGCCCAACGGCTGATCGAACGAAATGCCAATGCTGTTGCGCAGCTTGTGCTGGTCTGCG 5339
TGCCCAAGGGCTGATCGAAAGAAATCCCAATGCTATTGCGCAACTTGTGCTAGTCTGTG 14140
* * * * *

pVEIS01
374_J1a12 CGAAGCTGCGCAGGTGGGTTATGCAGTCTCAGAAAGCGGCTCGAACGGCTTACGTTTCCT 5399
CGAAGCTGCGTAAGTGGGTAATACAGTCTCAGAAAGCGGTTCAAAAGGCCTACGTTTCCT 14200
* * * * *

pVEIS01
374_J1a12 TGACGTGAGCCATGAAGGCTTCTTCGTCTTCAAACAGCCAGGCCGACACGCATTTTCT 5459
TGACGTGAGCCATGAAGGTGCTCTCATCTTCAAACAACCGAGGCTGACACGCATTTTCT 14260
* * * * *

pVEIS01
374_J1a12 TGATCGTGGCCATGTTGAACTGCCGTGCAGCTTCGACGTGTCCATGATGTAGAGCGCCA 5519
TGATCGTAGCTACGTTGAAGCTGCTGTGCAGCTTAGACGTCTCCATGATGTAAAGCGCCA 14320
* * * * *

pVEIS01
374_J1a12 GGCCGAGGAACAGGCTTTCAGCCGTGGAGGCCACACGTCATCACTTGGCGGGTACAGGA 5579
GACCAAGAAACAAGCTTTCAGCCGTGGAGGCCAAACATCATCACTGGCGGGTACAGGA 14380
* * * * *

pVEIS01
374_J1a12 TGTGTTATGCCCATCAAGTCGCTGATGCGAAAGCGGGTCATCGCGCACGTAGGCCA 5639
TATTGGTGATGCTCATTAAGTCGCTGATGCGAAACGCGGGTCATTGCGCACGTAGGCCA 14440
* * * * *

pVEIS01
374_J1a12 GCGGGTTCCAGCAGGCAGTTTGAAGTTCTCGTCATCAGGCGCGAAGTTATGGACTTCCT 5699
ATGGGTTCCAGCAGGCAGTTTGAAGTTTTCGTCATCCGCGCGAAGTTGTGAACCTCCT 14500
* * * * *

pVEIS01
374_J1a12 GCCCGCACTTCGCACGGAAGCCGAAGTAATCTCGAAGTTCTCGCCCTTGATGTCCAGCA 5759
GACCGCACTTTCACGGAAGCCGAAGTAATCTCGAAGTTCTCGCCCTTGATATCCAGCA 14560
* * * * *

pVEIS01
374_J1a12 CCACCACCGAATCCCGTAGTGGAGCAGGTTGGGAATCACGATGCCAACGCCCTTGCCGG 5819
CGACCACCGAGTCACGGTAGTGGAGCAGGTTGGGAATCACGATGCCAACGCCCTTACCCG 14620
* * * * *

pVEIS01
374_J1a12 AGCGTGTGCGCGCAGCCAGGAACATGAATTGCTGGCCACGGAACAGCAAAAAATCCTTGT 5879
AGCGTGTAGGTGCGGCCAGGAACATGAATTGCTGGCCACGGAAGAACAGAAAAATCCTTGT 14680
* * * * *

pVEIS01
374_J1a12 CTGCTTGCCCAACACGATGGAGGGGTACTTGTCTGCTCCGCTGGTGGGTTGAAGCAACC 5939
CCTGTTTTCCTACAAGAATGGATGGGTATTGATCATCAGGTTTGTGAGATTGTAGTAAC 14740
* * * * *

pVEIS01
374_J1a12 CACCCTTGACCAGCTCTTGCCGCGTGGCAACCGCGCCGAGCCGTGCAACTCCCGCTTTT 5999
CACTCTTGACCAGCTCTTGCCGCGTGGCGAACCCTGCGGAGCCGTGTAACCTCCCGCTTCT 14800
* * * * *

pVEIS01
TTGGGGTAAGCGCCGCTGATGCCGACCAGGCCGAACAGCAGCACCGGCAGGCCGATGCTGG 6059

```

374_J1a12      TTGGGGCCAGTGCAGCATAACCAATGATGCCGAACAGCATGACAGGCAGGCCAACGCTAA 14860
*****  ** ** ** ** ** * * * * * ***** ** * * * * *

pVEIS01        CCACCAGCGCGATAACGCCAAAGCCCTTTTCCTTCTTCGCGCCGAACCTTCGCCAGCTCGA 6119
374_J1a12      CCACCATGGCGATGATGCCGAACCCCTTTTCCTTCTTCGTGCCAAACTTGGCCAGATCCA 14920
*****  ***** * ** * * ***** ** * * * * *

pVEIS01        AGGGCAAGGACAGATGCGGGGTACGTCGCCCAGGCTCAGGTACTTCACCAGCAGGCCCG 6179
374_J1a12      AAGGCAAGGACAGATGCGGTGTAGATCCCAAGGTTAGATATTCACCAATAGCCCTG 14980
* ***** ** * * * * * * * * * *

pVEIS01        AGGTCACGTAGAGCATCCCAAAGACGGATGCCACGGCCAACAGCAGCAGCACCAGGACGG 6239
374_J1a12      AACTCATATAGAGCATGACAAGAACAGAGACCATGGCCAATAACAGCAGCACCATAACGG 15040
* *** ***** ** * * * * * ***** * *****

pVEIS01        CGATGGCTATTTTCTTCTGCTCATGACGCGCGGCTTCGAGCAGTTGGAGTTTCTTCAC 6299
374_J1a12      CGATGGCTATTTTCTTCTGTTTATGATGACATGCCCCGCAACAGTTGAAGTTTCTTAC 15100
*****  ***** * * * * * * * * * *

pVEIS01        GGGGTCAAAGTACAGCTCTTTGAGGTGGGTTTCTCGAAGAAAACCACCACATCAATGGT 6359
374_J1a12      GGGGTCAAAGTACAGTTCTTTAGGTAGTTTTCGAAGAAAACAACCACGTGATGGT 15160
*****  ***** ***** ***** ***** ** *****

pVEIS01        CGTGACCACCTCTTGCATGATGTAGTCGAAATCCAGCGTTTGCCGACTTCCGATTGCTT 6419
374_J1a12      CGTTATCACCTCTTGCATGATGTAGTTAAATCCAACGTTTGCCGACTTCTGATTGCTT 15220
*** * ***** ***** ***** *****

pVEIS01        GATGTACGAGCCGATTTTGTAGAGCGCACCCCGCAATCGTTGCGGTGGATAGAGGTCAC 6479
374_J1a12      GATGTAAGAGCCAATTTGTAAAGAGCACACGGCAATCGTTGCGGTGAATGGAAGTGAC 15280
*****  ***** ** ***** ***** ** * * * * *

pVEIS01        AGAGCCGCTATGCCCGGACTTCAACGCCATAAGGTAGTCCCATGTTTCATCCCCGCGCAG 6539
374_J1a12      GGAACCGCTATGCCCTGACTTCAAGGCCATGAGGTAATCCCATGTTTCATCACACGCGAG 15340
** ***** ***** ***** ***** *****

pVEIS01        CTCAGTGATGAACAGGTGATCGGGCTTCATTTCGCATGCAGCTCGACAACACTGCTTTGGC 6599
374_J1a12      CTCGGTGATAAAAATATGATCTGGTTTCATGCGCATGCAGCTCGATAACACTGCTTTGGC 15400
*** ***** * * ***** * ***** *****

pVEIS01        GGTGATGGTCGGGCTGAAAAACAGGTGTACCGAGTTCGATGCTTGGGCGTGGTCAGCTC 6659
374_J1a12      GGTGATGGCTGGACTGAAAAAAGGTGTACCGAGTTAGGATGCTTGGGTGTAGTTAACTC 15460
*****  ** ***** ***** ***** ***** ** * * * * *

pVEIS01        GTGGGCATCCTCAATCGTGAACAGGCGCGCAGGCCGGGTAGAGGTCAATCAGGCACTT 6719
374_J1a12      GTGAGCATCCTCAATCGTGAAGGCGGCACGAGGTCGGATATAAATCAATCAGGCACTT 15520
*** ***** ***** ***** ***** * *****

pVEIS01        GGAAACGTCGTTTTCGCGGAGCCGTCGCCGCCACGGTGACGATGTTCAAGTGATGCTC 6779
374_J1a12      TGAGAACGTCGTTTTCCTGACCTGTGCCGCCACCGTCACGATGTTCAAGCGATGTT 15580
** ***** * * * * * ***** ** ***** *****

pVEIS01        TACCGCCAGCTCAAAGAAGCGCACAAACTCGCCTTGTGGGCCAACCGGAGCAATTCAGC 6839
374_J1a12      TACCGCTAGCTCAAAAAGCGCACGAAATCGCCCTTTTGAACCAATTGAAGCAATTCAGC 15640
*****  ***** ***** ** ***** * * * * *

pVEIS01        CTCCCAGGGCTGAATTTCTTCCGACAAGCCAGCCAACACGGGCTTGAGCCGCCCGGAATC 6899
374_J1a12      CTCCAAGGCTGAATCTCATCCGAGAGCCCGCCAACACTGGCCTCAGTCGTCCAGAATC 15700
*****  ***** * * * * * * * * * *

pVEIS01        CGCGTAGTTGGCCAACGAAAACCGGCTGGTCGAAGGCTTGCGAATCGTCATGCTCACGGT 6959
374_J1a12      CGCATAGTCTGTAGTGAAGAACGACTGCTTGAAGGCTTGCGAATCGTCATGCTGACCGT 15760
*** ***** * * * * * ***** ***** *****

pVEIS01        GCCGTCTTCGACGCGCGGAGGCAACACGATTTGGCAACGCTCCCCATCGGGCAGCGTCAG 7019
374_J1a12      GCCATTTTCGACGCGAGGAGGAGTACAACCTTGGCAGCGTCCCCATCAGGGAGCGTGAG 15820
*** * ***** ** ***** * * * * * * * * * *

pVEIS01        CGAGGCAATGGGGTTGTGCGCGTGCAGCTTCTTTGGTTGAACACGGCTATGGCCGTGGC 7079
374_J1a12      TGAGGCAATCGGGTTGTCTCTGTGCGAGCTTCTTCTGGTTGAACACGGCTGTAGCCGTGGC 15880
*****  ***** * ***** ***** ***** * *****

pVEIS01        CAACCCCTGACAGTTTCGACAGCGACAACGATTATTGGGCACCTCACGCCAGCCCGCCGA 7139
374_J1a12      CAACCCCTGAAGGTGAGAGAGCGACAGTACTCGTTAGGCACCTCACGCCAACCCGCGAGC 15940
*****  ***** * * ***** * * * * * *****

pVEIS01        ATCCTCAATCCAGTAGCTGCCTGGCCTGTTGATAGCAAACCTCGGTGATGTTTGCCATCAA 7199
374_J1a12      GTCCTCAACCCAGTAACCTGCCTGGCCTGTTGATCGCAAACCTGTGATGTTTGCCATCAA 16000

```

```

*****
pVEIS01      GGCTTCGGAGATACCGAATTTCGCGCATCAGCTCCCGCACCGAAAGGAGCGGTCAACCCG 7259
374_J1a12    TGCTTCGGAGATACCGTATTTCCGCGCATCAGCTCCCGCACAGACAGGGATCGGTCAACCTG 16060
*****

pVEIS01      GCTTATCGGCTTGAGCACAGCCTCACTCAACAACGTTAAAGACATTTTCAAAGCTCACGT 7319
374_J1a12    ACTCACCTGTTTTTGTGGGACATCATTTCAACAAGGTTAAAGACATTCTCAAAACTCACGT 16120
*****

pVEIS01      CACGCGCCACCATGACGTTTACAATCGTCGCGGAAGCACCACAAAGTGGGTGGGATGT 7379
374_J1a12    CACGCGCCACCATGACGTTTACAGCCGTCGCGAGGAGTACCACCAGAGTGGGAGGGATGT 16180
*****

pVEIS01      TGATCGAATTGCGCAACGCCTCAGCGGCCATATCTTGGGTGGCCTGATTGGTGTTCGAT 7439
374_J1a12    TGATAGAGTTGCGCAATGCCTCCGCTCCCATGTCTGGGTGGCCTGTGTGGTGTTCGAT 16240
*****

pVEIS01      ACGGCTGTGCGAGCCGTCGCGCCGTGTGCTGCTATTGCGCCAGCGCTGCGAGTAGGCTT 7499
374_J1a12    ACGGTTGAATAACCGTTGTACCGGTATTGCTTCTATTGCTACACGCTGGGAATAGGCTT 16300
*****

pVEIS01      GTGAAAAGTCCTTAATCAGGCTCAGCATGATTGCGCCGCCGAACCGCTGCCAAAAGTGGG 7559
374_J1a12    GCGAGAAATCCTTAATCAGACTCAGCATGATTGCACTGCCGAACCGCTGCCAAAAGTGGG 16360
*****

pVEIS01      TGTCACAGTAGCCGGGAATACCGCTGTAGCCCATTTGGTCAGTCGCCGGGCTGTCGATGT 7619
374_J1a12    TATCTACGTAGCCAGGAATACCGCTGTAAACCATTTGATCGGTGCGTGGGCTATCGATGT 16420
*****

pVEIS01      TGGCAAACACCCCGTTGGGGTTGTCGATGCGCGTCCAGACAACGAAGTTTCGTTCTTGTC 7679
374_J1a12    TGGCAAACACCCCATTAGGATTATCAATGCGCGTCCACACAATAAAGGTACGTTCTTGAC 16480
*****

pVEIS01      CTGTGCTTGAGCTGCGATTGCTGCTCACCGAACCGTCTGCCCGCGCTCGACCAACAACA 7739
374_J1a12    CTTGCTGGAGCTGCGATTGCTGTTGCGCAAAGACCGTTGCGCGCGCTTCAATTAACAACA 16540
*****

pVEIS01      CTTTGCCATTGGCCGAGTACACGTTGTGAGCACCTTGCAAATGACGAAGCCGGGCAGCG 7799
374_J1a12    CTTTGCCATTGGCCGAGTATACGTTGTGAGCACGTTGCAAATGACAAAGCCTGGCAGCG 16600
*****

pVEIS01      TCGTATCAATGCCGTTTTCAGCGCGCAAGGAATCGTGGTGC CGCGCTTCAACAGGTAGT 7859
374_J1a12    TCGTATCAATGCCAGTTTTCAGCGCACAGGAATCGTGTGCGCGCGCTTCAACAGATAAT 16660
*****

pVEIS01      CTAGGTTGGGCAGCTTGCCCGCCAGCGGGCTTGCAAGACCGTAGGTTGCAGCCGGGCCG 7919
374_J1a12    CAAGATTAGGCAGCTTTGCAGCGCAAACGGGCTTGCAAGACCGTAGGTTTCAGACGCCCG 16720
*****

pVEIS01      CAAGCCCATCAGCGCCCCCGCCCTGGCCACCAGGCGCGAGCTGCAGGCCAGTTCACCAT 7979
374_J1a12    CAAGCCCATCAGTGCCTTCACTTTGGCCAACAGGCGTGAGCTGCTGGACACCGCCTTGAT 16780
*****

pVEIS01      TGAAGCCC-TGGCCACCGCCCGGAGTCAGCCAGCCGCGCCATACGGGCTTGCACTTCG 8038
374_J1a12    TCATGCCCATCGCTGCCATGCGGGCCTGGACTTCGCGTTGCTGC-----TGTCTCTTC 16833
*****

pVEIS01      CGTTGCTGCTGTTCTTTTCGACAACTGCGCATCAGCCTTGGGCTTCGCGCCCTTGCTG 8098
374_J1a12    TCTGAAGAAGACGTATTTCTGACAACGGCGCGTCGGTTTTGGCTTTGCGCCTGTGTTC 16893
*****

pVEIS01      TCCAGGAGCACGCCGCCGCCATCTTGCCTTCCATCGGCGTGGGTGGGGAGGTGGCTGG 8158
374_J1a12    TCCAGGAGCACACTGCCGCCATCTTCCGCTCCATCGGCGTAGGTGGGCGTGGTGGTTGA 16953
*****

pVEIS01      TTGCCCTGCTGCGCCTGGCCACCGCTGTACTGCCCGCGCGTTGTTGCCCGTCAAGCCTGG 8218
374_J1a12    TTTCTTGTCTGC-----TGCACT-TGAGCAGAATCG----CCTTGAGCGCGT 16995
*****

pVEIS01      CCAAACCGCCTTGGCCGCGTTTTCGCGCCTCGGCTGCGGCCTGGGCCTTCGCTTGCGCC 8278
374_J1a12    TTGCA--GCATTGGCAGCA-----GCCAAGGTCTCAGCTTGGATTCTCGCTTGCGCC 17046
*****

pVEIS01      TGAGCCTCTTCTACGGCCTTGATTGGGCTTGATCTTCTCTTCTTCTTCTTGATTTCG 8338
374_J1a12    TGAGCTTCTTCTATCGCCTTGATCTGAGCTTGTCTTCTCTTCTTCTTCTTGATTTCG 17106
*****

```

pVEIS01 GCCTTTTGCCCTTCTCAATTAAATCACTCTCGACCGCCGATTCTTCGACGCAAATTCAGGC 8398
 374_J1a12 GCCTTTCTCTTTTCAATCAAATCACTTTTCGAGCGATGCATTATTGAAGCCAACCTCAGGT 17166
 ***** ** * ***** ***** ** * ***** ** * * * *****

pVEIS01 TTCACCGGCGCAGTCTTGGCCGGTGTCTCCACGGTTGCATTGCCCCGCTTGTGCTTTTGG 8458
 374_J1a12 TTCAATGAATCACTCTTAGCGGGTGTCTTCACAGATGGATGAGCCAGCTTGTACTTCTGG 17226
 ***** * ** * ***** ** * * * ***** ** * * *

pVEIS01 TAGAACAAAAGCCCGGCCACGATAAACAGGAACCCACAAGCACCAGGCCAAGGATGGCG 8518
 374_J1a12 TAGAACAAAAGCCTCCACCATGAACAAAACGCCATAAGCACCACCAAGGACGAGG 17286
 ***** ***** ** * ***** ** * ***** ***** * *

pVEIS01 AATTTGCCCACATTCTGTTTTGGACTTTGCTTTGCTGGTCTGGCTCGCCCAAACCTCGGGT 8578
 374_J1a12 AATTTGCTTACATTTGTTTTGACTTCGCCTTGTGCTGGTCTGGCTCGCCAAAACCTCGGTT 17346
 ***** ***** ***** ***** ** ***** ***** ***** *

pVEIS01 TCAGTGTCCAACGTGGGCAATGCGCCACGTTTGTGAAATCGGTAGTCTCGCTCATGGTT 8638
 374_J1a12 CTTCATCCAGCGTGGTAATGCCCGCGTTTGAGAAATCAATCGTTTATTATCATGATT 17406
 ***** ***** ***** ** ***** ***** * * * ***** *

pVEIS01 GGGTGCCCTTATCTACACGCACAGCGCCCGGAACGCTGGTGCCGGTTTTCGTTGAATTTGG 8698
 374_J1a12 AGATGTCCCTATCTGCACGCACGGCGCCAGGAACGCTGGTGCCGGTTTTCGTTGAATTTGG 17466
 * * * ***** ***** ***** ***** ***** ***** *

pVEIS01 GTAGCGTGTACGCCCGGTTAATAAATTTCAATGACCTGATCGCCCATCCGTGCCCGCACGG 8758
 374_J1a12 GAAGTTTGTACGCACGGTTAATAAATTTCAATGACCTGATCGCCCATAGGCGCGCCCTTACGG 17526
 * * * ***** ***** ***** ***** * * * ***** *

pVEIS01 TGCGAATCACCTCATGCAGAACCCAGGTTTGTATTGTTGCTCAGGGTCCACGTTGCTATTCA 8818
 374_J1a12 TACGAATCACCTCATGCAGTACCACGGTTTGTATTGTTGCTCAGGATCAATGTTGCTATTCA 17586
 * ***** ***** ***** ** ***** ** * *****

pVEIS01 ACAAGCGCTCAGTGCCATCAGGCAGAACCTTGAAGAACACCGGCAACTCGCCGGAGTGGT 8878
 374_J1a12 ACAGTGCTTCTGTACCATCCGGGAGCACCTTGAAGAACACCGGAAGCTCGCCAGAGTGGT 17646
 *** ** * * ***** ** * ***** ***** ***** *****

pVEIS01 CATACACCAGGCGCGTGAATCGCCCATCGTCATAAGCGGCTGTGGGCTTCAACAACCTCGT 8938
 374_J1a12 CATAGACAAGGCGGTGAATCGCCCATCGTCGTAAGCGGCGGTGGGCTTCAATAACCTCGT 17706
 ***** * ***** ** ***** ***** ***** ***** *****

pVEIS01 TATCACCGCGCCAAACGTAATCCGTGTTGATTACCGTTTCTCAGCTCGCGCAGCGGCGG 8998
 374_J1a12 TATCACCGCGCAGAGTAATCTGTATTGATGATCATTTTTCGCGCGGGGTACGGCAG 17766
 ***** ***** * ***** ** ***** * * ***** ** * * ***** *

pVEIS01 TAATCTGCACCCGTTTGGCTTCGGCTGCCGCTCGGCGCAGCCTTGGCCGCTTCGGTAT 9058
 374_J1a12 TTGTCTGTACTCGTTTGTCTTCGGCAGCTGCTT-----TCGAGGCTTCTGTAT 17814
 * ***** ** * ***** ***** ** * * ***** * * ***** *****

pVEIS01 CCAGATACCGGAAACGAGAGGGCGAAATACCAAGGATTGAGTCTTGGCCGTAGCCTGCAACT 9118
 374_J1a12 CTGGATACCGGAATCGCAGAAATACGTGGGCATTGAATCCTTGGCTGTAGCCAGCAATT 17874
 * ***** ***** ***** ***** ***** ***** ***** *****

pVEIS01 CGAATGAATACGTGCGCTTATTCGTACACGACGACGTTTGTGTTGTGGCATGTTCTGCG 9178
 374_J1a12 CAAACGAATATGTGCGTTTGTGTTGCTACTACAATCACATTGTTTGGCGCATCTTCTGCG 17934
 * ** ***** ***** ** * ***** ** * ***** ***** ***** *****

pVEIS01 TGGGCTTCAAAACAATGTTGTTGCCGCGCAGCCCAAATTCGAAGCGTCAGCGTCACCAA 9238
 374_J1a12 TTGGCTTTAAACAATATTGTTGCCACGTACCCCAAGGTCCCAAGCCTTGGCGTCACCAA 17994
 * ***** ***** ***** ** * * * ***** * ***** *****

pVEIS01 TGCCCGCAGCACGAAGAGGGCGAAATACCAAGGATTGCTCTTCTCGAGCTGGATTAGGG 9298
 374_J1a12 TACCCAACAATGAAGAGGGAGAAATCTCCAAGGATTGCTCAGCTTCGAGCTGGATGAGCG 18054
 * ***** ** ***** ***** ***** ***** ***** ***** *****

pVEIS01 AAGCGCGGCCAATTTGAGTGTGCACGTTGTACACCACGACTTGTGCTAAATCTCAGTGC 9358
 374_J1a12 AAGCACGGCCAATTTGAGTGTGTAAGTTGTAGACAGAAGACTTGTGCTAAATTTCACTAC 18114
 ***** ***** ***** ***** ***** ***** ***** ***** *

pVEIS01 GAATGCGCGGGTCAATCCGCATGGGCCAAACTGGCGCACAGGCCAGCACCAGGGCGATTA 9418
 374_J1a12 GAATACGCGGGTCATTGCGATACGCAAAATGAGTAACACAAGCCAGTAGCAAGACAGGTA 18174
 ***** ***** ***** ** * * * ***** ***** * * * * *

pVEIS01 GTTTTTTCATTTAATCACCTCGGGGTCAACGCGGTAAGTTGCGGCCTTGAATCCGAGCGG 9478
 374_J1a12 GTTTTTTCATTGAATCACTTCGGGATCAACCGGTAAGTTGCAACCTTAAATCCAGGGG 18234
 ***** ***** ***** ***** ***** ***** ***** ***** *

pVEIS01
374_J1a12
ATTCACAAGCCGTTGTTGCTCAGTCATCATGCCGGGCTTGAATTGGAATGCAATGGTGGC 9538
ATTGATCAGTCTTTGCTGCTCAGTCATCAAACAGGTTTGAATTGAAAAGCGATTGTGGC 18294
*** * ** * *** ***** ** * * * * * * * * * *

pVEIS01
374_J1a12
AATCCACTTCTGCAGGGGCGAGTTATCGACGTTCTCACCGCTCGCACTCAGCTTTTCCGA 9598
AATCCACTTTCTCACTGGAGAATTATCGAGATTCTCACCGCTAGCACTGAGTTTTCAGT 18354
***** ** * * * * * * * * * * * * * * * *

pVEIS01
374_J1a12
GGTGAAACGCACCTTGGGCGGTTTACCGAAAAACGCAATGGAGGTGATTTTCACAACCAC 9658
GGTGAACCGTACTTGGGCCGTATCACCGAGAAAGGCAATAGAAACGATTCCAACCACTAT 18414
***** ** * * * * * * * * * * * * * * * *

pVEIS01
374_J1a12
CTTGCCCTTTGTCTTCAACACCGACAGGGGCGAGCCAGGCGCTTGCACCTTGC CGGAATA 9718
TTTTCCCTTGTAGCAAGGACTTTTAGCGGTGAGCCTGGGGCATCCACTTTGCTGTCATA 18474
** *

pVEIS01
374_J1a12
TTCCTTGGCCACATCGCCATCAGACATGAGCTTGCACGCCTCGAATTGCTCAGAAATGGT 9778
CTCTTTGGCCACATCGCCATCTGACATGAATTGACGCAATTGAACGATGATCGTAACTGT 18534
** *

pVEIS01
374_J1a12
GAACCAGTCGTAACCTTCACAGGTGCGGACGTATTGGGCCAACCAATACTTGTTTACCAC 9838
GAACCAGTCGTAACCTTCGCAAGTACGACGATGAGCTAACCAATACTTGTGACGAC 18594
***** * * * * * * * * * * * * * * * * * * *

pVEIS01
374_J1a12
CTCGTCGTATTGGTCTTTGGTGTGCGCGCACCGAGCGCAAAACAGTTGTTGCGCCCGTGCT 9898
TTCATCTAGTGTGCTTTGGTGTGACGACGGAACGCAAAACAGTTGTCGCACTGTACT 18654
** *

pVEIS01
374_J1a12
CTTGTC AACCTGCAAAATCGTGGGTT CAGGTTCCGTGCGCGTCAGCAGTGCGACCAGGAA 9958
CTTGTC AACCTGTAATAATCGTGGTTCGTTT CAGAGCGCGTCAGCAGAGCTACCAGGAA 18714
***** * * * * * * * * * * * * * * * * * * *

pVEIS01
374_J1a12
CGCAAGAATAGACACGCTGCAAAATGGCCGTGGACACCATGGCAATAATCCACGCTATTTT 10018
CGCCAGAAATAGATACGCTGCAAAATCACCGTAGCAACACTTGCGATGATCCAGGCAATCTT 18774
*** *

pVEIS01
374_J1a12
GCGGGAGCGCTTAATTTCTGAATCTTGGACTTCTCAAAATCTGCCGCGAGTTCCGCAAAA 10078
GCGGTGAGCGCTTAATCTCTTCAATCTTGATTTCTCGAAATCTACCGCGCTTCCACGAA 18834
*** *

pVEIS01
374_J1a12
GTTTTGTGGAACCGGGCAGGCTTTGCGGCCCTTTTGGGCGTAGGTTTCTCTTTGATGGT 10138
AGTTTGTGGAAGACGGGCGAGCTTATTGAGC---ACAGGTGC---TTGCTCTTTGTACT 18888
***** * * * * * * * * * * * * * * * * * *

pVEIS01
374_J1a12
CATGGTTTAGCCTGCTTTCTGATCGAATTGAATTTGTTTGCCGCCCAGCGGCCAACTGC 10198
CATAATTTAGCCTGCTTTCTGATCGAATTAAATTTGTTCTCTGCCCATCGACCAGCAGC 18948
*** *

pVEIS01
374_J1a12
TTGGCCATCGGCTTTGGTGGCGGATAACCGGCCTGCATGGCCTTTATGTCCTTTGCGGA 10258
CTTGCCAGCATTGATGGCGGTACGCCCTCTTGCTGGATGCCCTTTGCTCCTAACCTGGC 19008
* *

pVEIS01
374_J1a12
GCCTTTGATGGCACTTACCAGCGTGCCTATGCTCCCCGGCATTGAGGCATTGATACCGCC 10318
TGCCTGAGCCGCATGTCTCAGCGTATTAGCGCTTCCAGGCATTGATACATTGATGCCGCC 19068
* *

pVEIS01
374_J1a12
CGCGCTCAGGGACGACGCCAGGGCACTCACCTGTTTCAAAGGAATAGCAGCGTCAGATT 10378
AGCACTCAGCGATGACGCCATGGCACCACCTGTTTCAAGAGAAACAGCAACGTCAAATT 19128
* *

pVEIS01
374_J1a12
GCCCATTGATGCAAGGAACACCGCTTTCAACGACACGTCATCAAGCGAGCCACTGCCGGA 10438
GCCCATGGATGCCAGAAATACCGCTTTAAAAGAAATATCATTCAATGTGCCATTGGCGGA 19188
***** * * * * * * * * * * * * * * * * * *

pVEIS01
374_J1a12
AAGCATGGTTTTTGAAAGACCCACGGATATGGTGGCAATCACCGCCACGAACACGGGAAT 10498
AAACATTTGTTTTGAAAGTCTTACGGATATGGTGGCAATTACCGCCACAAATATAGGCAT 19248
** *

pVEIS01
374_J1a12
CAGGGCATAGCTCAACGCTGTGTTGAGCCAGGCCGAGAAATATTGCCGTGTCGCGGGAAA 10558
CAGTGCATAGCTGAATGCCGTATTGAGCCATGATGAGAAATACTGACGCGTCGCTGGAAA 19308
*** *

pVEIS01
374_J1a12
TATCAGGAAGGCGAAAAACAGCGGCCCGATCATGGCGACCAAGACAGAGCCAACATCAGC 10618
TATCAAAAGGCAAGTACATTGGACCCACCATAGCCACCATCACAGAGCCAACATCGGC 19368
***** * * * * * * * * * * * * * * * * * *

pVEIS01 AATGATGATTGCAATCGTTGCTGCAACCAGGAACGGAATCAGGCCGACCAGGATGCATGT 10678
374_J1a12 AACGATGAGGCAAAGCGTTGCGGCAACAAGGAATGGAATCAGGCCAACCAGAATGAAGAT 19428
** ***** ** ***** ***** ***** ***** ***** ***** *

pVEIS01 GTATTTTCATGAAATACAGGAGCATGGAACCCGCATTCCATGGAATCCGCTGCTGCTGGC 10738
374_J1a12 AGCTTTGAGTCCATAGATCAACATTGCCCCACACTGCCAGGAACTTGAGGGCTTTTAC 19488
*** * *** * * * * * * * * * * * * * * * * *

pVEIS01 CGATTTCGTAGCCATCATCAAGAATCTTCAGGTAGTGGAGGGCGAGTTGGTCTAACGTGCC 10798
374_J1a12 AGATTGATAGCTCTCATCAAGAATATTGAAATAGTGCAGCGCCAGTTGATCAAGCGTGCC 19548
**** ***** ***** ***** * * ***** * * ***** * * *****

pVEIS01 AGCGGAGGCGCTGCCACCGCTGACCGCGGAAGCCAAATCACTGCCTACGCCAGTCACGAT 10858
374_J1a12 AGCATTTGGCGTGCCACCGCTAATCGCTGATGCTAAATCACTGCCTACACCGTCACGAG 19608
*** ***** ***** * * * * * * * * * * * * * * * * *

pVEIS01 AGGAATAACCGTGCTGGTGTAGTTTCGCGGCATTCAAGCCAGGCCAATAATCACAGCGAA 10918
374_J1a12 AGGAATCACCAGGCTGGTGTAGTTAGCGGCATTCAAACCGAGGCCAATAATCACCAGCAA 19668
***** ** ***** ***** ***** * * ***** ***** * * *

pVEIS01 ACCCGCGCACCGAAGCGCAAAATCAATTACAGGCTCAGTCTCTGCGCCGCGCATGTAATT 10978
374_J1a12 GCCCGCGCACCGAATGCCAAATCAATCACCAGGCTCAGTTTCTGCGCCGCGCATGTAATT 19728
***** ***** ***** * * ***** ***** ***** *****

pVEIS01 GATGCAGATCAGGATGATGTAAATACCGAAGCCAAACCGCAGCCAGGGCGGTTACCGCCGT 11038
374_J1a12 GAGGCAGATCAGAAGGATGTAAATCCCAAAGCAAATGGATGCCAGCGCGGTGACCGTTGT 19788
* * ***** * * ***** * * * * * * * * * * * * * * * *

pVEIS01 GGCAATGGCTGATGCGGCGGAACAGCGCCAGAACTCACAATCTTGTTTATTTACCGAC 11098
374_J1a12 AGCAATGGCTGACGCGGCGGCATCAGCGCCAGAACTCACAATCTTGTTGATTTCATCAAT 19848
***** ***** * ***** ***** ***** ***** ***** *

pVEIS01 AAGCCAAGTAATTGGCGCTGCTTCATCCATGGTATTATCTCTTGAACCTTGTTGTTATCTA 11158
374_J1a12 GAGCCAGGTAATTGGCGCTGATTTCGTCCATGG-GTTACCTTTTATACCTTTGCGGTCCA 19907
***** ***** ***** * * ***** * * * * * * * * * * * *

pVEIS01 AACAGGCATGGTCTGTCCCGCAGTGGTTGGAGGCTTGCTAGCCAGCACTTTGATAACG 11218
374_J1a12 AGCAATCGTGGTCCGTTGTGCGAGTGGATATGGGTTTAGTTGGCCAGCACTCTTGTA-- 19965
* *

pVEIS01 GGCGCGAGAAGGCTTTAGCCGCGAAATAAGCAGACCAGCGGCAGCAGCGTTCTGCACTC 11278
374_J1a12 -----ACCTTCAAGCTGGACAACAGCAC--CAGTCGCATTAGTCCATC 20006
* *

pVEIS01 GGGCGTCAGATTGATACTTGGCTGGGTGGACCTGTGCATGTGAGAAATTTTCTAGTCT 11338
374_J1a12 TTGGATCATTTTCGTACTTTGAGGGTCAGTCAAGCTAGCTTATTCATAGCCCTAGCCG 20066
* *

pVEIS01 CGTCCATGTCTGTCATGTAATCGGCCACCGAAGGGACTTTCTCACTTGAGCAACCGGCAA 11398
374_J1a12 CGTCTATATCGTGCAGTACTCAGAGGTAGACGGCACCTTTTCTCGACTGCAAGCAGTGA 20126
***** *

pVEIS01 CGATTAAATGGCGCTGACGATCAGCAGCCTGAGCGCTTGAATCTTTGTCATGTGTTCTC 11458
374_J1a12 TGACCAACACGGTGCAGACTATTAAACAACCGGAACGTTGAATCTTTGTCATGACAACCT 20186
* *

pVEIS01 CTT---GGTTGCCTGTAATGCAGCTTCTCAATCTCTTGACGGTGCTTTTGTGGAAGAAT 11515
374_J1a12 CTCTACGGCGGCTTGGCTTGCATTTCC---AACTCCTGCTTTGCTTAGCTTGGAGAAAT 20243
* *

pVEIS01 GGTCAACAGATTTTGGTTAGCCTGAATAGCGTTTGTCTCGTGATAAGGCGGTTAGCCAA 11575
374_J1a12 GGTCAACAGGTTTGGTTGCTTGAATGGCGTTTGTTCGTTGATCAGGCGATTGGCCAA 20303
***** *

pVEIS01 ATCGGCTTTAGCAGCCGGGTGCTTGGCTGAATCAATCTGCGCATTACGGGACTGAATAAG 11635
374_J1a12 GTCAGCTTTGCGCGCAGGATCATTCGCGGAATCAATCTGTGCTGTCAGCGATTGAATCAA 20363
* *

pVEIS01 GCTCAATCGGTTATTGGCCTTGCCGTATAAGTCGCCCATCACTGCGTTCTGCGAAGCAAT 11695
374_J1a12 ATTCAATCGGCGATTGGCCTTGTCTGTATAAATCACTCATGATCACGTTGAGACGCAAT 20423
***** ***** ***** * * * * * * * * * * * *

pVEIS01 CACGTCATACAGTGCATTGCTTTGGGCAAGCGCTCAATGTGCGGTATTTCTTGCGCTC 11755
374_J1a12 CACGTCATACATAGCATTTGCTTTTGAGATTCCCGAGGTCGGATATTTGCTGCGCTC 20483
***** ***** * * * * * * * * * * * * * * * *

pVEIS01 ATTGATGTAGGCACTGGTCTTCTTTACGTCAGTGAACCTGTTGCCAGTCAGCGGGCAA 11815

```

374_J1a12      AGTGGCATAGGCGCTGGTTTTCTTCACGTCACCTCAGTACCTGTTCCAGTCAGAAGGCAA 20543
* * * * *

pVEIS01        TGCAGCTTTTCAGCGACGAGTTATCCAGCAAAGCACCCATGCCACGCGGGCCAGTGATGGC 11875
374_J1a12      CGAGGCTTTTCAGTGCCGAGTTATCCAGCAAAGCGCCCATGCCACGCGAGCCAGTAATGGC 20603
* * * * *

pVEIS01        CGCGTACTGTTGCTGATACTGTCTGAATTTGTTGCTGCAATTGGCTGAACTGCAGGCCCA 11935
374_J1a12      TTCGTATTGCTTTGATGTTGATCAATTTGCTGCTTTAACTGCTCGAACTGCTGCTTCCA 20663
* * * * *

pVEIS01        CTTGGCAATGCTTTTCGACCTGGTTTTTAATGCTGGTCGAAATGTGGGCAACGTCAGTCAC 11995
374_J1a12      TTTCGCAATGCTTTTCGACATGTTCTGAAATGCTGTCGAAATGTGGCAGCGTCATAAAC 20723
* * * * *

pVEIS01        TGAATCTGCGCTTGCAGCAGCTCCAGCGAAGACGACAACCAAAGCGCCACTGGCCGCTCT 12055
374_J1a12      CTGGAGCTGCGCGTATGCAGTTCCAGCGAAGATAAACACAAAGCACCTTATAGGCACCT 20783
* * * * *

pVEIS01        GGTTGTCTTCAAAATACCTTTTCAGATCCATGGTCATTCTTCACGGTTGCGGGTTGAGAA 12115
374_J1a12      GGTTTTCTTCAAAATTCCTGTAAGTCAATTTCTTCAGGTTACGGGTTGAGTA 20843
* * * * *

pVEIS01        ACTTTTTTCGCGGCTTTTTTTCGCGAGCACCTTGTAGAGGGGCACCCACTGGTCAGGATGC 12175
374_J1a12      ACGTTTTTCACGCGCTTTTTTACGGAGTGCTTATAGACAGGCACCCACTGATCAGGATGC 20903
* * * * *

pVEIS01        GTGCCAGCTTGCCTTGGCAGCATCCAGATACTTGAAGTCTTCGGCATCCATTGCCAGC 12235
374_J1a12      GTGCCCGCTTGCCTTGGCCGCTCCAGATACTTGAATCCGAGCATCCATTGCCAGT 20963
* * * * *

pVEIS01        ACGGCAATGTGCATACCCATGCGGTTCAAATCCAGCTTGGCCAAGCTGGATTGCGAACCT 12295
374_J1a12      ACGGCAATGTGCATCAGCCATGCCACTCAAATTCAGCTTGACCAGGCTGGATTGCGAACCC 21023
* * * * *

pVEIS01        TGCTTGACCATAAACATGCGGCTTTGCAGCCCAAGTTCTTGAGCTTTTGAAACTCTTTG 12355
374_J1a12      TGCTTGACAATGAACATGCGGCTTTGTAGTCCGAGTTCTTCAACTCTGAAACTCTTTA 21083
* * * * *

pVEIS01        GTGGTCAAGCCGAAGCGCAATAAACCCGCGCTGCGGTGGTGTATTCCGCGTCAGGG 12415
374_J1a12      A--TCTGAGTAAACGGAATAACCGCCACCGTCAGGCGTGGTGTATTCCGCGTCAGGA 21140
* * * * *

pVEIS01        TTGGGCAAAATAAATCTTGGTCGGTGTCTGTTGCAGCACCGCAGGCAGCAGCGGCTTCTA 12475
374_J1a12      TTTGGCAAAAAATCTTGGTGGGTGTCTGCTGTAGCACCGCAGGAAGTAGCGGCTTTTA 21200
* * * * *

pVEIS01        ATCAGCTCCTCGGGGACTGCGACACCAACAATGAATTCACCCGCGCAGCGCTGTT 12535
374_J1a12      ATGACATCCTCGGGGACTGTGAAACCAACAATAAATTCACCCGACGACACCGCTT 21260
* * * * *

pVEIS01        TTCAGGGAATCAAGAATCTGGTCTGCCGTGGTGGGTATTGCAACGGCAACCAGAACTCT 12595
374_J1a12      TTCAATGAATCAAGAATCTGATCTGCCGTGTGCGGTATTGCAGAGGTAACCAGAACTCT 21320
* * * * *

pVEIS01        TCAACGACCGTGGCCAGCAACGCGCCGTCGCGCTGCATCAAGGTCTTGAGGTGGAACAGG 12655
374_J1a12      TCAACTACCGTGGCGAGCAACTCGCCATCACGCTGCATCAAGGTCTTGAGATGGAAGAGA 21380
* * * * *

pVEIS01        TAAGCCAAAATTGGCTCAGTGGCTGGGTGGCCAGCCACCAAGAAATCCGACACGTCAAAG 12715
374_J1a12      TAAGCTAAAATTGGCTCAGTTGCAGGATGGCCAGCCACCAGAAAATCAGATACATCAAAG 21440
* * * * *

pVEIS01        CCAACGCGCTTGAACCTTCCCAATCGAAGTTGTTTTGCGGGTTGTCCAGGGCGTAGGCA 12775
374_J1a12      CCAACGCGCTTTAAATCTCCCAATCAAAGTTATTTTGTGGATTATCCAAGGCATAGGCA 21500
* * * * *

pVEIS01        TAACGTCCGTCCACCTCGCCATAGCACCAAGTCAGCCAGGCGGCGGCAAGCAGTTTTTCG 12835
374_J1a12      TAACGTCTCTTATTTTACCATAGCACCAAGTCAGCCAGGCGGCGAGCCAAACAGTTTTTCG 21560
* * * * *

pVEIS01        CCACGGTCGGGGATGCTTTGCAGAAGCACCCGAAGCGCCGTTGCTCAAACGGCAGCTCA 12895
374_J1a12      CCATGATCAGGGATGCTTTGTAATAGAAGCCAAAGCGCCGTAGCTTGAAGGGCATCGCA 21620
* * * * *

pVEIS01        AACACGTGTGTCACCGCCCGTTTAATGTCTTGGTGTCTCTGCGGTGCCCTCTTGGCCA 12955
374_J1a12      AACACATTGTCTACGGCTGCTTAATGTCTTGGTGTCTCTGCGGTGCTCTCTTGGCCA 21680

```

```

*****
pVEIS01      GCCTTGGCGGCCACACGCGCCACCAGGTCATAGAGGAAGTTGCGGTTCTGCGGTGTGTCTG 13015
374_J1a12    ACTTTGGCGGCCACACGCAACCACCAGGTCATAGAGAAAGTTGCGGTTTTCGCGCGTATCT 21740
*  *****  *

pVEIS01      GGCAGTTGGAAGGGATTCAAGCCCGTAGGCTCGCCGCTGGCCAGCCGGAAGTACGTCCCG 13075
374_J1a12    GGAAGTTGGAAGGGGTTTAAACCCGTAGGTTCCGCACTGGCCAGCCGGAAGTACGTCCCT 21800
**  *****  *

pVEIS01      CCCACGCTTTCGATGAAGCCGCGCATAGAGCCGTCCTTGTCCACGGCAAACAGCTTGTGTG 13135
374_J1a12    CCCACGCTTTCGATAAAGCCGCGCATAGAGCCGTCCTTATCCACGGCAAATAGCTTGTTA 21860
*****

pVEIS01      TCCAGCCTTCCCAGGAAGGTCAAGAGCGTGGTTTGCAGCGTGGTCTTACCGGCACCAGTC 13195
374_J1a12    CCCCATCTGCTACAGAAGGTCAATGCAACAGTTTGAACCGTAGTTTTACCAACACCCGTC 21920
*  *  *  *****  *****  *

pVEIS01      GCGCCCAAAATGACCGTGTGGCCAGCGCGTTTTCGCGCGTGAATCCAGCTCAGGCAGC 13255
374_J1a12    GCGCCAGTTATTAAAAGGTGGCCAGCCACCTTTTCACCGTGAGAGTCCAGATCGGGCGGC 21980
*****  *  *  *****  *****  *

pVEIS01      GAGTAGTGAAAGTTGAAGTGGTACACGCCATGCACGGTTGTTTGCAGCGGCATGACCGCC 13315
374_J1a12    GAGTAGTGAAAGTTTAAAGTGGTACACACCATGCGCCGGTGTTCGAGCGGCATGATCGCT 22040
*****

pVEIS01      GAACCGTCGCCCCAAAGGATTGCGGTATGCTTGCCCGAGCTGTAGGTGTTTATGCTGAAC 13375
374_J1a12    GAACCATCACCCACAGGATTGCCATATTGCTTGCCCTGAATTGTATGTGTTTATGCTAAAC 22100
*****  *

pVEIS01      AAACCCAACAAATTCCGGGTTGTTTGGGCATCGGACGCGGCGGCGTTGACGTTGCC 13435
374_J1a12    AACCCAAGAAAATTCTGGTCGTTTGGGCATCGGACGCGGACGCGGCGTTGACGTTTCCT 22160
*  *  *  *****  *****

pVEIS01      GGGAACATCGAGAAGAAGCTTTCCGGCGCTGAAATGGTGGCGGCGACAACTGAGTCGCG 13495
374_J1a12    GGAAACATCGAAAAGAAGCTTTAGGTGCTGAAATAGTGGCAGGAACGAACCGGGTTGCA 22220
*  *****

pVEIS01      CATGAACCACTGAGCGTAGTCCGGGCGCTCGCGCCCCGGTCTTCGGTTTGCTTTTCGGTC 13555
374_J1a12    CATGAACCACTGAGCGAAGCTAGAGCCGTAGCGCCCCGATTTTCGGCCTGCTTTTCAGTC 22280
*****

pVEIS01      TTACCATACACGGCCAACGCGCCGTGCAGTTCACCGAAATACACCTGGCCGGACATGATC 13615
374_J1a12    TTTCCGTACACGGCCAACGCGCCATGCAGTTCACCGAAATACACTTCGCCAGCCATGATG 22340
*  *  *****

pVEIS01      GCGCCTTTGCCAGGCTCATATCGTCCATTTGCTCCACCGCCTCATCACCAGCCGATTCC 13675
374_J1a12    GCACCTTTGCCAAAGTCCATATCTTTTATCTGCTCAGTGGCTTCGTCGCCCGCCGATTCC 22400
*  *  *****

pVEIS01      ATTTTGTTCAGCGACTGGTTAATCATTCGGATGGAGTCAGCCGTATTATGAAGGTGAAC 13735
374_J1a12    ATTTTATTCAGGCGCTGATTATCAACCTGATAGAGTCAGCCGTATTGATGAATGTAAAT 22460
*****

pVEIS01      GACAAGCACATGATGAAGGGAAACGGCAGCTCCAGCAGCGGTTGAATTTCCCCCGGTC 13795
374_J1a12    GACAAGCACATGACAAAAGGAAACGGTAACTCTAACAGGGGGTTGAATTTCCCCGTGTT 22520
*****

pVEIS01      GTGGGATCCGGGAAGTCCTTCAAATCGAAAAACGCGCTGTAGCGGTTGCCACCGTCCGGG 13855
374_J1a12    GTTGGTCCGGAAAAATCCTTTAAATCGAAAAACGCACTATAGCGATTTCCACCATTTGGG 22580
*  *  *****

pVEIS01      AACCGCGTTTCAAGCAGCTTGTAGCCATGGTGCAAAGCGCTTGTTCGACTGCCTGAAAC 13915
374_J1a12    AACCGCGTTTCTAGAAGCTTGTAAACCATGGTGCAAAGCACTTGTTCGACCACCTGGAAC 22640
*****

pVEIS01      AGCGGGAGCGAGGTCACGGGGACGCGCTCCCAAAGCCGTTGTACAGGTACGCAATGAAC 13975
374_J1a12    AGCGGAAGAGAGGTGACAGGAACGCGCTCCCAAAGCCGTTGTACAGGTACGCAATGAAC 22700
*****

pVEIS01      TCGTAGAACTGACTGAACTGGTTGCCATCGTGCTCGTACATCGACAGCACTTCGCATTCA 14035
374_J1a12    TCGTAAACTGACTGAACTGGTGGCCGTATGATCGTATATCGACAACACTTCGCATTCA 22760
*****

pVEIS01      TAGGGCGCAACGGTTTGCCTCACACTGAGCTGAATTTCTCCAGCTCACGAATGCACCTCT 14095
374_J1a12    TAAGATGCAAGGTTTGTGTCAGATGATCTGCATCTTCAAGCTCACGAATGCACCTCT 22820
*  *  *****

```

pVEIS01 TCCAGGTTGTTCATTCATGCCCCGGTTTCAGGATGAAGGTCAGATAAAAAATCGTTCTCGAAA 14155
374_J1a12 TCAAGAGAGTCATTCATCCCTGCGTTCAGAATGAACGTCAGATAAAAAATCGTTCTCGAAA 22880
* * * * *

pVEIS01 ATATCTGTGCCTTCAAACCTTGGCCATGTACTTGGCACTGAAAAGGCGCATCCATTCGTAG 14215
374_J1a12 ATAGATTCACCTTCAAACCTAGCCATGTACTTCTCTGTGAAAAGTCGAATCCATTCATAG 22940
* * * * *

pVEIS01 CCAAATCGGTAGTCCGTTTTGAACTCCGTTTGGTAGTGGTCGAGGTGTGCCACACTGCC 14275
374_J1a12 CTAAATTCGTAATTTGATTTAAGCGTCTTTGCGTAGTGGTCGATGTGCGCCCCACCGCT 23000
* * * * *

pVEIS01 AGGCGCGAGCCGGTGGACTTGGCTATGCTCAGGAGCAGCCCGTTCAACCCATCGTATTGG 14335
374_J1a12 AAACGTAGCCCTGTTGACTTGGCCAAGTCTCAGAAACAGGCCCGTTCAACTCATCGTATTGG 23060
* * * * *

pVEIS01 TTCTCCAATACCTTGTCTGCTGGTCACTTCAAAGGGCAGGCTCTTTCCTTGAGAATGAAC 14395
374_J1a12 TTCTCTAATACATTTGTCGCTCGTCATTTCAAAAGGCCAAACCTTTGCTTGAGAATGAAT 23120
* * * * *

pVEIS01 GCCACGCGCCCATCCTCAAGAGACACAATCGTGTCTCTGACGTGGTAGTTATACGGGGGC 14455
374_J1a12 GCCATAGCCCATCTTCAAGAGAAACGATCGTGTCTCTGACATGGTAGTTATAGGTTGCT 23180
* * * * *

pVEIS01 AGTTGTTCTCTACCCGCGAACTGTTTTTTTAAATAGCGCTTAACATTTTGTAGCTTGCG 14515
374_J1a12 AGTTGTTCTTACCCGCGAACTGTTTTTTTAAATAGCGCTTAACATTTTGTAGTTTACG 23240
* * * * *

pVEIS01 ACCGTAGGTTGTTGGGGCAATGGCCATAGTCCCGCCGTGATATTTGGCGTTGCCACTGAG 14575
374_J1a12 ACCGTAAAGTTGTTGGAGCAATGCCATTGTTCCGCCGTGATATTTAGCGTTGCCATTGAG 23300
* * * * *

pVEIS01 CGCTTTAAGAATGCTCCATTTTACTTCGAGCATCAGAATTTGGATTGCCTTGTCTGTCATT 14635
374_J1a12 CGCTTTAAGAAGGCTCCATTTCACTTCGAGCATCAGAATCCGAATGGCTTTATCGTCATT 23360
* * * * *

pVEIS01 GGCGCAAAGCATCTTGACGAACAAGGCAACTGGCACCCCGATAAGGGCAAACAGCCAGCC 14695
374_J1a12 TGTGCAATGATCTTGACGAACAATCATTGGCACACCAATAAGAGAAAACAGCCAGCC 23420
* * * * *

pVEIS01 CAGGCCACCAACGAAGGTGCCGAGCAATAAACCGCCGATCAGCGATAAACACATGATCGC 14755
374_J1a12 CCAGCTAGCAACAAGGTACCAAGCAATAAACCGCAAGCAGTGATAAACACATGATCGC 23480
* * * * *

pVEIS01 CAGGCCGACCATGTACGAATGCCCCAGATCATTGGGCTACGCCAAGACCGTTGTAGCT 14815
374_J1a12 TAATCTGCCATGTACGAATTCGCCAGATCATTGGGCTACGTCCAAGACCGTTGTAGCT 23540
* * * * *

pVEIS01 GATGTACTTTTCTGGTGGCTGTTCGTCTGTCTATGGCATCAGGCGAACAGTGTCAGGCC 14875
374_J1a12 GATGTACTCTTCGGGTGATTGTTCGTCTGTCTATGGCATCAAGCGAACAGTGTCAGGCC 23600
* * * * *

pVEIS01 AGGAGCCGAGCGCAACAGCGCCGCCACCAGAGAGACGTACACCACGCCCCACCCAAAAT 14935
374_J1a12 AGCCACCAAGCGCCACCGCACCACCGACCAATGAGACGTACACCACGCCCCACCCAAAAT 23660
* * * * *

pVEIS01 CTGCCCCAAGATTTTCTCGGTGAATGCCATCACACCGAGATAAATCAGGTAAGTCAGGG 14995
374_J1a12 CTGCCCATGATTTCTTCTCAGTGAATGCCATCACACCGAGATAAATCAGGTAATCATGG 23720
* * * * *

pVEIS01 CGACGACACCGACTAGGGCATAAAGCCCGGTCTTAATGTCGTCAGCAGCGGTTTTCGCCT 15055
374_J1a12 CAATGACACCAACCAGGGCATAAAGTCTGTTTTATGTTCTCAGCGGCTGTTTTGCCCT 23780
* * * * *

pVEIS01 TATCCAGGCCACCCAAT-----GCGAGAGCGTTGGACGACGCGGCAGCAATCG 15103
374_J1a12 TTGTGAGACCACCCAATCCTCCAACCTGACGCGAGAGCATTTGACGACGTAATAGAAATCG 23840
* * * * *

pVEIS01 CCACAGCCGAGCCACCTTACGGTGTGGGCAATCAGCGGGTTGAAAATGCGGTTTCATGG 15163
374_J1a12 CAATAGCTAAAATTACATTATGATATTTATCTAAAAAATATTAAAAATATGTTTCATAT 23900
* * * * *

pVEIS01 CTAATTTCTTTACGGCCAAAGTTTTTAAAAATATCCCAACCCCTCGTAAGTGACTTTTTTC 15223
374_J1a12 CTAATTTCTTTAAGACCAAGTTTTTAAAAATATCCCATCTTTTGTAGTGATTTCTTC 23960
* * * * *

Material Suplementar

Pirosequenciamento do DNA genômico de *X. fastidiosa*

1. Sequenciamento de DNA em 4 etapas com o *GS FLX Titanium System* (454 Life Sciences/Roche) realizado no CATG

Etapla 1: construção da biblioteca de ssDNA.

A amostra de DNA genômico deve ser transformada em uma biblioteca de fragmentos de DNA fita-simples (ssDNA) adequada para o sequenciamento com o Sistema FLX 454. O método para fragmentar o DNA genômico (nebulização ou sonicação) varia de acordo com o tipo de amostra e com o objetivo do experimento. Aos fragmentos obtidos são ligados adaptadores que possibilitam a imobilização dos fragmentos em microesferas.

Etapla 2: amplificação clonal baseada na realização de PCR em emulsão (emPCR)

Moléculas de DNA fita simples obtidas da biblioteca de fragmentos serão imobilizadas em microesferas (*beads*) de forma que cada uma dessas esferas não transporte mais de uma molécula de DNA amplificável. As *beads* contendo moléculas de DNA serão emulsificadas com os reagentes para amplificação do DNA por PCR em uma mistura de água e óleo, de modo que cada *bead* com uma única molécula de DNA ligado seja capturada em uma micro-gota suspensa no óleo, constituindo assim um microreator de amplificação. A amplificação da biblioteca é realizada simultaneamente, sendo que os fragmentos amplificados permanecem ligados a mesma microesfera original em cada microreator, resultando em *beads* cobertas com ~10 milhões de cópias de um único fragmento de DNA.

Etapla 3: sequenciamento no 454 *Genome Sequencer FLX* (454 Life Sciences/Roche)

Após a amplificação, as *beads* recobertas com fragmentos de DNA são depositadas na placa de sequenciamento (PicoTiterPlate) que contém 1,6 milhões de poços; cada poço tem a dimensão exata para comportar uma única microesfera. A placa é então inserida no equipamento, e os reagentes do pirosequenciamento são sucessivamente adicionados aos poços da placa. O sequenciador FLX executa automaticamente e monitora a sequência de reações em todos os poços da placa simultaneamente.

Etapla 4: processamento e análise de dados

Após ~10-12 horas de corrida, são obtidos dados brutos das imagens dos sinais luminosos gerados a partir da PicoTiterPlate e detectados por uma câmera CCD. Após ~8

horas adicionais de análise computacional, as imagens são convertidas em sequências, cada uma derivada das imagens de um poço da placa.

As diferentes etapas envolvidas no sequenciamento estão esquematizadas na Figura 1, e foram realizadas em salas distintas do CATG do IQUSP, as quais foram preparadas segundo as recomendações da 454 Life Sciences/Roche. O procedimento utilizado no sequenciamento dos genomas de *X. fastidiosa* seguiu rigorosamente os manuais da 454 Life Sciences/Roche e foi realizado com Felipe Beckedorff e Murilo Sena, orientandos do Prof. Sergio Verjovski-Almeida, coordenador do CATG.

2. Escolha do tamanho adequado da biblioteca de DNA genômico e do número de regiões que serão ocupadas na placa de sequenciamento

O tamanho da biblioteca de DNA genômico e o número de regiões da placa 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. No caso do sequenciamento de genomas considerados pequenos (como o de *X. fastidiosa*) ou se o interesse é obter apenas cobertura parcial ou de baixa profundidade da biblioteca de DNA, opta-se pelo sequenciamento paralelo de bibliotecas distintas utilizando-se 4, 8 ou 16 regiões da PicoTiterPlate (PTP), as quais são delimitadas pela colocação de uma máscara (*gasket*). Por outro lado, o sequenciamento de genomas maiores que exigem maior profundidade de cobertura, deve-se utilizar pelo menos uma placa inteira (2 regiões) carregada com uma única biblioteca. A Tabela 1 mostra o número de regiões na placa e o rendimento esperado para um sequenciamento gerando *reads* com tamanho médio de 350 a 450 nucleotídeos, segundo informações do fabricante. Desta forma para sequenciar um genoma bacteriano de ~4 Mbp com cobertura de 8 vezes (32 Mbp de dados a serem processados) poderia ser utilizadas apenas uma região de uma placa dividida em quatro regiões. Nós optamos por carregar as regiões 1, 3 e 4 da PTP com bibliotecas de ssDNA distintas de (Tabela 2), as quais correspondiam ao DNA genômico extraído de *X. fastidiosa* cepas J1a12 (15 e 25 passagens em cultivo) e 9a5c (14 passagens em cultivo). A região 2 foi carregada com uma biblioteca de ssDNA preparada de uma amostra de DNA genômico bacteriano submetida a um ensaio teste que incluía uma pré-amplificação por PCR dos fragmentos nebulizados de DNA visando enriquecimento da amostra.

Figura 1. Resumo das etapas com o Sistema GS FLX Titanium System (454 Life Sciences/Roche).

Tabela 1. Rendimento do sequenciamento com o *GS FLX Titanium System*^a

Tamanho da placa (PTP)	Regiões ocupadas na placa	Tamanho da região	Número de <i>beads</i> da biblioteca aplicadas por região	Número de <i>Reads</i> esperado /região (x10 ³)	Bases/região (Mbp)
70 x 75 mm	2	30 x 60 mm	2.000.000 (x2)	450-650	180-280
70 x 75 mm	4	14 x 43 mm	790.000 (x4)	160-250	60-110
70 x 75 mm	8	5,4 x 60 mm	340.000 (x8)	80-120	30-55
70 x 75 mm	16	2 x 53 mm	125.000 (x16)	25-40	25-40

^aTabela adaptada do manual do *GS FLX Titanium System*

Tabela 2. Bibliotecas de ssDNA carregadas na PTP

Região	Biblioteca de DNA genômico
1	J1a12 com 15 passagens (J1a12 15P)
2	DNA genômico nebulizado submetido à pré-amplificação
3	J1a12 com 24 passagens (J1a12 24P)
4	9a5c com 14 passagens (9a5c 14P)

3. Reagentes utilizados

Foram utilizados os reagentes e kits da série titânio: *GS FLX Titanium General Library Preparation*, *GS FLX Titanium emPCR*, *GS FLX Titanium emPCR Breaking*, *GS Sequencing* e *GS Titanium PicoTiterPlate* (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 7500 LabChip kit* (Agilent), *RNA Pico 6000 LabChip* (Agilent) e *Quan-iT RiboGreen* (Invitrogen).

4. Construção da biblioteca de DNA genômico

4.1. Fragmentação do DNA e recuperação dos fragmentos

Amostras de 4 a 5 µg DNA genômico de *X. fastidiosa* (9a5c com 14 passagens em cultivo e J1a12 com 15 e 24 passagens em cultivo) purificado e quantificado como descrito

em Procedimentos Experimentais (item 3.9.) foram fragmentadas por nebulização para obtenção de fragmentos com tamanho de 500 a 800 bp. A nebulização foi realizada em fluxo laminar para evitar a dispersão de DNA no ambiente de trabalho, utilizando-se aparato que é parte do *GS FLX Titanium General Library Preparation Kit* (454 Life Sciences/Roche). O volume da solução de DNA preparada em TE 1x foi aumentado para 100µL pela adição de TE 1x. Esta solução foi homogeneizada com a adição de 500µL de tampão de nebulização e transferida para o aparato de nebulização. Para a quebra do DNA foi aplicada pressão de 2,1 bar de nitrogênio, por 1 min. Apenas a solução que permaneceu na base do copo de nebulização foi coletada, o que correspondeu a um volume aproximado de 390µL para 9a5c 14P e 370µL para J1a12 15P e J1a12 24P. A estes volumes foram adicionados 2,5mL de tampão PBI (*MinElute PCR Purification Kit*, Qiagen) para a homogeneização e posterior filtração em colunas MinElute. 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. Para garantir que 80% dos fragmentos recuperados estivessem nessa faixa, a quantidade adequada de *beads* a ser utilizada foi previamente calibrada, sendo que utilizamos a proporção de 0,7 *beads* por molécula de DNA. A qualidade e a faixa de tamanho dos fragmentos da biblioteca dupla-fita de DNA genômico foi verificada com o *DNA 7500 LabChip kit* no BioAnalyzer (Agilent) (Figura 2).

A

Figura 2. Continua na próxima página

B **C** **D**

Figura 2. Análise das bibliotecas de DNA dupla-fita das cepas de *X. fastidiosa* no BioAnalyzer (Agilent). Em **A**, está representada a separação das 3 bibliotecas de fragmentos de DNA dupla fita por eletroforese em gel capilar utilizando o *DNA 7500 LabChip kit* (BioAnalyzer, Agilent). *Sample 1*, *2* e *3* correspondem, respectivamente, ao DNA genômico de J1a12 15P, J1a12 24P e 9a5c 14P nebulizado e purificado. O marcador de tamanho em pb está mostrado na primeira canaleta (Ladder). Pode ser observado um enriquecimento nos fragmentos na faixa de 300pb e 1000 pb. Em **B**, **C** e **D** estão representados gráficos de cada amostra separadamente, onde o eixo y mostra as unidades de fluorescência e o eixo x, o tamanho em pares de bases (pb). Os dois picos maiores, nas regiões de 50pb e 10380pb, correspondem à fluorescência medida para o marcador de tamanho. O perfil de distribuição dos fragmentos de DNA gerados pela nebulização de cada uma das amostras de DNA genômico está mostrado entre os dois picos maiores.

4.2. Preparação da biblioteca de ssDNA

Os fragmentos de DNA dupla-fita obtidos como descrito acima foram reparados com T4 DNA polimerase e ligados aos adaptadores A e B. Estes adaptadores são um par de oligonucleotídeos dupla-fita que possibilitam a imobilização dos fragmentos de DNA nas microesferas (Figura 3), 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 as *reads*. O adaptador B possui biotina na sua extremidade 5' que garante a imobilização da biblioteca de DNA dupla-fita nas *beads* de imobilização (que possuem estreptavidina acoplada) e o subsequente isolamento da biblioteca de DNA fita-simples 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.

Após a imobilização dos fragmentos de DNA das bibliotecas dupla-fita as *beads* magnéticas ligadas a estreptavidina, segue-se preenchimento de falhas deixadas após a reação de ligação dos adaptadores utilizando-se a enzima *Fill-In Polymerase* (GS FLX Titanium General Library Preparation Kit).

Figura 3. Imobilização dos fragmentos de DNA dupla-fita nas microesferas. Quatro tipos de produtos são gerados pela ligação dos adaptadores A e B aos fragmentos de DNA dupla-fita: DNA ligado aos adaptadores A e B (produtos AB e BA) que correspondem a 50% dos produtos de ligação, DNA ligado a adaptadores B nas duas extremidades (produto BB) e DNA ligado a adaptadores A nas duas extremidades (produto AA). Os produtos AA não se ligam às microesferas, uma vez que os adaptadores A não possuem biotina (representada pela esfera em cor laranja). Esquema extraído do manual do GS FLX Titanium System (454 Life Sciences/Roche).

O passo seguinte foi o isolamento de bibliotecas de DNA fita-simples pela desnaturação (com NaOH) da dupla-fita dos fragmentos imobilizados nas *beads*, e consequente liberação do DNA fita-simples com a extremidade não biotinizada. O DNA fita-simples (biblioteca de ssDNA) é recuperado no sobrenadante, após separação das *beads* utilizando o *Magnetic Particle Concentrator* (MPC).

A qualidade das bibliotecas de ssDNA foi determinada utilizando-se o *RNA Pico 6000 LabChip* no BioAnalyzer (Agilent). A Figura 4 mostra que a nebulização das amostras de DNA genômico de 9a5c 14P, J1a12 15P e J1a12 24P gerou conjuntos de fragmentos de ssDNA com os tamanhos médios de 650, 610 e 620 nucleotídeos, respectivamente. A qualidade e a faixa de tamanho dos fragmentos das bibliotecas de ssDNA foram verificadas com o *DNA 7500 LabChip* no BioAnalyzer (Agilent) (Figura 4).

A quantificação das bibliotecas de ssDNA foi feita com o *Quant-iT™ RiboGreen® RNA Assay Kit* (Invitrogen). O *Quant-iT RiboGreen* é um reagente não fluorescente quando em solução. No entanto, quando ligado ao RNA ou a DNA fita simples emite fluorescência em $530 \pm 12,5\text{nm}$ quando excitado em $485 \pm 10\text{nm}$. Foram preparadas duas curvas padrão, *low* e *high-range*, para garantir a determinação correta das concentrações das bibliotecas de ssDNA. A curva padrão *low-range* é indicada para determinar concentrações entre 1ng/mL e 50ng/mL (Tabela 3). Já a curva padrão *high-range* é indicada para determinar concentrações entre 20ng/mL e 1µg/mL (Tabela 4). Em ambas as curvas o RNA utilizado como padrão foi uma amostra de RNA ribossomal de *E. coli*. A concentração das bibliotecas fita simples também foi então determinada com base nestas curvas padrão, sendo que os valores de unidade de absorbância obtidos para cada amostra foram multiplicados coeficiente angular da reta, extraído da curva padrão mais adequada (Tabela 5).

Figura 4. Análise das bibliotecas de DNA fita-simples das cepas de *X. fastidiosa* no BioAnalyzer. Em A, está representada a separação das 3 bibliotecas de fragmentos de ssDNA por eletroforese em gel capilar utilizando o *RNA Pico 6000 LabChip* no BioAnalyzer (Agilent). *Sample 1, 2 e 3* correspondem, respectivamente, as bibliotecas de ssDNA de J1a12 15P, J1a12 24P e 9a5c 14P. O marcador de tamanho está mostrado na primeira canaleta (*Ladder*). Pode ser observado um enriquecimento nos fragmentos na faixa de 600 nucleotídeos. Em B, C e D estão os gráficos de cada amostra separadamente, onde o eixo y mostra as unidades de fluorescência e o eixo x, o tamanho em nucleotídeos (nt). O pico maior na região de 25 nt corresponde à fluorescência medida para o marcador de tamanho, seguindo-se o perfil de distribuição dos fragmentos de ssDNA.

Tabela 3. Preparo da curva padrão *low-range* para quantificação das bibliotecas de DNA fita simples com o *Quant-iT RiboGreen* (Invitrogen)

Volume de TE 1x (µL)	Volume de RNA padrão (100ng/mL) (µL)	Volume de <i>Quant-iT</i> <i>PicoGreen</i> 1x (µL)	Concentração Final do padrão <i>Quant-iT</i> <i>PicoGreen</i>
0	1000	1000	50ng/mL
500	500	1000	25ng/mL
900	100	1000	5ng/mL
980	20	1000	1ng/mL
1000	0	1000	Branco

Tabela 4. Preparo da curva padrão *high-range* para quantificação das bibliotecas de DNA fita simples com o *Quant-iT RiboGreen* (Invitrogen).

Volume de TE 1x (µL)	Volume de RNA padrão (2µg/mL) (µL)	Volume de <i>Quant-iT</i> <i>PicoGreen</i> 1x (µL)	Concentração Final do padrão <i>Quant-iT</i> <i>PicoGreen</i>
0	1000	1000	1µg/mL
500	500	1000	500ng/mL
900	100	1000	100ng/mL
980	20	1000	20ng/mL
1000	0	1000	Branco

Figura 5 – Quantificação das bibliotecas de DNA fita simples das cepas de *Xylella fastidiosa* com *RiboGreen* (Invitrogen). No gráfico, o eixo y representa a intensidade de fluorescência do sinal emitido em cada diluição das amostras da curva padrão de dosagem de ssDNA fita simples e o eixo x representa a concentração de RNA padrão do *Quant-iT RiboGreenRNA Assay Kit* (Invitrogen) nas diluições que utilizadas para determinação da curva padrão *low-range*).

Tabela 5. Concentração das bibliotecas de ssDNA das cepas de *X. fastidiosa*

Biblioteca	Sinal de Fluorescência	Concentração ng/μL
J1a12 15P	2,96	3,30
J1a12 24P	2,13	2,49
9a5c 14P	1,63	1,70

5. PCR em emulsão (emPCR)

Uma vez determinada a concentração de cada biblioteca, é necessário definir a concentração de moléculas de DNA presentes em cada uma delas. Para tal, é utilizada a equação: Número de moléculas de DNA)/μL = concentração da biblioteca fita-simples x $6,02 \times 10^{23}$ / 328,3 x média de tamanho dos fragmentos de DNA. Onde $6,02 \times 10^{23}$ é o número de Avogadro e 328,3 corresponde a média de peso molecular dos nucleotídeos em g/mol. A média de tamanho dos fragmentos de DNA foi obtida através da análise do perfil da

biblioteca de DNA dupla-fita com o *DNA 7500 LabChip* no BioAnalyzer (Agilent). O resultado obtido para cada biblioteca de ssDNA está apresentado na Tabela 6, sendo então utilizado para preparação de diluições seriadas de cada biblioteca de 1×10^8 a 1×10^6 para cada biblioteca. Essas diluições foram estocadas a -20°C até realização da emPCR.

Tabela 6. Número de moléculas de DNA/ μL para cada uma das bibliotecas de ssDNA das cepas de *X. fastidiosa*

Biblioteca	Concentração pg/ μL	Tamanho médio dos fragmentos de DNA (pb)	Moléculas de DNA/ μL
9a5c 14P	1700	650	$4,8 \times 10^9$
J1a12 15P	3300	610	$9,92 \times 10^9$
J1a12 24P	2490	620	$7,37 \times 10^9$

A reação de emPCR requer que cada *bead* esteja ligada a apenas uma molécula molde de ssDNA. 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 DNA 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. Para otimizar a quantidade de cada biblioteca a ser usada na emPCR nós optamos pelo método de titulação da emulsão. Através desse método, é possível determinar a porcentagem de *beads* enriquecidas e através desse número, estimar o resultado do sequenciamento. Para tal, volumes diferentes de cada biblioteca na concentração de 1×10^6 moléculas de DNA/ μL foram misturados a um número fixo de *capture beads* (Tabela 7).

Tabela 7. Proporções entre DNA fita simples e *capture beads*.

Proporção	Volume da solução de DNA (1×10^6 moléculas/ μ L)	Volume da solução de <i>beads</i> ($2,4 \times 10^6$ <i>beads</i> / μ L)
0,5:1,0	1,2 μ L	1 μ L
1,0:1,0	2,4 μ L	1 μ L
2,0:1,0	4,8 μ L	1 μ L
4,0:1,0	9,6 μ L	1 μ L

Os fragmentos que constituem a biblioteca de ssDNA devem ser imobilizados nas *capture beads*, para permitir sua separação na emulsão. O objetivo nesta etapa 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 em condições controladas, para formação da emulsão. Estas reações foram preparadas utilizando o *GS FLX Titanium Small Volume emPCR Kit*, seguindo-se emPCR com o seguinte programa de amplificação: 94°C por 4 min, 50 ciclos de 94°C por 30 seg, 58°C por 4,5 min e 68°C por 30 seg, sendo mantidas ao final a 10°C até serem retiradas do termociclador. As emulsões foram então quebradas para a recuperação e lavagem das *beads* contendo a biblioteca de ssDNA amplificada.

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 *beads* da biblioteca de ssDNA efetivamente amplificada. Esta etapa baseia-se no anelamento do iniciador de enriquecimento, o qual é biotinilado, ao adaptador A de cada molde amplificado, e a posterior ligação desse iniciador 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 iniciador de enriquecimento. A proporção final entre a quantidade *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.

Desta forma a quantidade ótima de biblioteca de DNA a ser utilizada na emPCR definitiva (que será utilizada efetivamente na etapa de sequenciamento) é aquela que resulta na fração de 8% de *beads* enriquecidas (uma faixa de 6 a 10% é aceitável) e corresponde à quantidade que resultou em uma única molécula de DNA utilizada como molde para a amplificação (Tabela 8).

$$\% \text{ de } beads \text{ enriquecidas} = \frac{\text{Número de } beads \text{ enriquecidas}}{2,4 \times 10^6 \text{ beads}} \times 100$$

$$2,4 \times 10^6 \text{ beads}$$

Para determinação dessa porcentagem foram interpolados os resultados mostrados na tabela 8 (Figura 6). A porcentagem de *beads* recuperadas e enriquecidas foi satisfatória para 2 das bibliotecas de ssDNA, uma vez que a porcentagem mínima de *beads* recuperadas na titulação deveria ser de 65% e a de *beads* enriquecidas deveria estar entre 5% e 15%. Entretanto, a porcentagem total de *beads* enriquecidas da biblioteca de J1a12 15P ficou abaixo dessa faixa e por isso foi necessário definir a proporção de cópias de DNA por *bead* para emPCR definitiva a partir da extrapolação dos resultados obtidos na titulação dessa biblioteca, de forma que esse número se aproximasse ao máximo da porcentagem ideal de 8%. A extrapolação definiu o número de 8 cópias de DNA por *bead* (Figura 8A), já que com 4 cópias de DNA por *bead* a porcentagem de enriquecimento foi de apenas 5,17%. Para as bibliotecas de ssDNA de J1a12 24P e de 9a5c 14P, as proporções foram de 2 e 1,75 cópias de DNA por *bead* respectivamente, correspondendo a 8% de *beads* enriquecidas. Com base nesses resultados foi realizada a amplificação definitiva por emPCR. A recuperação e o enriquecimento das *beads* finalmente obtido estão apresentados na Tabela 9.

Tabela 8. Titulação das bibliotecas de ssDNA de *X. fastidiosa*

Biblioteca fita simples	J1a12 15P			
Cópias de DNA por bead	0,5	1	2	4
Total de beads recuperadas	1670000	1773200	1859000	1180690
Total de beads recuperadas (%)	69,58	73,88	77,46	49,2
Total de beads enriquecidas	53754	84456	88580	71757
Total de beads enriquecidas (%)	2,24	3,52	3,69	2,99
Biblioteca fita simples	J1a12 24P			
Cópias de DNA por bead	0,5	1	2	4
Total de beads recuperadas	2121000	1750320	1933000	1889000
Total de beads recuperadas (%)	88,38	72,93	80,54	78,71
Total de beads enriquecidas	49140	89868	367500	418425
Total de beads enriquecidas (%)	2,05	3,74	15,31	17,43
Biblioteca fita simples	9a5c 14P			
Cópias de DNA por bead	0,5	1	2	4
Total de beads recuperadas	1944745	2344175	2157850	2143050
Total de beads recuperadas (%)	81,03	97,67	89,91	89,29
Total de beads enriquecidas	147500	164285	212160	450500
Total de beads enriquecidas (%)	6,15	6,85	8,84	18,77

Figura 6. Titulação das bibliotecas de DNA fita simples de *X. fastidiosa*. As bibliotecas de ssDNA de *X. fastidiosa* foram tituladas com o *GS FLX Titanium Small Volume emPCR Kit* para a determinação do número ideal de *beads* por cópias de ssDNA. Estão mostrados os gráficos referentes à J1a12 15P (A), J1a12 24P (B) e 9a5c (C). O eixo y representa a porcentagem total de *beads* enriquecidas e o eixo x, as proporções de cópias de DNA por *bead* utilizada na titulação. Em todos os gráficos, foi traçada uma reta para permitir a extrapolação dos dados e melhor definição do número ideal de cópias de DNA pelo total de *beads*.

Tabela 9. Resultado da emPCR para as bibliotecas de ssDNA de *X. fastidiosa*

Biblioteca	Cópias de DNA por bead	Total de beads recuperadas %	Total de beads enriquecidas %	Total de beads enriquecidas
J1a12 15P	8	80,93%	18,77	1.823.200
J1a12 24P	1,75	74,48%	5,25%	469.500
9a5c 14P	8	66,08%	15,11%	1.197.900

6. Pirosequenciamento

A etapa de sequenciamento é dividida em três fases: o preparo da placa PTP, a pré-lavagem do sequenciador e a corrida de sequenciamento, com esquematizado na Figura 7. É importante sincronizar o preparo 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. Se o instrumento estiver em funcionamento contínuo, procede-se o encerramento da corrida anterior, seguido pela substituição do cartucho de reagentes pelo cartucho contendo as soluções de lavagem.

Após a reação de emPCR, as *beads* ligadas à biblioteca de ssDNA são aplicadas na placa PTP, a qual foi previamente carregada com outras *beads* que garantem o empacotamento e a formação de camadas de apoio. Foram carregadas 4 regiões da PTP, 3 delas com o volume total obtido das bibliotecas de ssDNA de J1a12 15P, J1a12 24P e 9a5c 14P após emPCR e enriquecimento das *beads*. 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, uma única *bead* com moléculas de ssDNA idênticas resultantes da etapa de emPCR. A reação é iniciada a partir de um iniciador de sequenciamento que é anelado à 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 sulfúrilase 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 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. As imagens são processadas pelo *software* que acompanha o 454 *Genome Sequencer*, determinando a sequência dos fragmentos de DNA (*reads*) das bibliotecas. 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 *reads* processadas são então submetidas a filtros de qualidade que identificam e excluem aquelas que apresentam baixa qualidade. Finalmente as *reads* 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. Os dados brutos, metadados e dados métricos são armazenados em planilhas e arquivos texto, a partir dos quais informações sobre o sequenciamento podem ser geradas. 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*.

Figura 7 – Etapas do sequenciamento no 454 Genome Sequencer FLX.

Súmula Curricular

Patrícia Isabela Pessoa da Silva

Nascimento: 13/07/1976 - Nova Lima/MG - Brasil

Formação Acadêmica

- 1998 - 2002** Graduação em Ciências Biológicas.
Universidade Federal de Minas Gerais, UFMG, Belo Horizonte, Brasil
- 2004 -2010** Curso de Doutorado, 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

- 2004** Curso de curta duração em *New Approaches in Drug Design*.
Universidade de São Paulo, USP, São Paulo, Brasil
- 2007 - 2008** Programa de Aperfeiçoamento de Ensino.
Universidade de São Paulo, USP, São Paulo, Brasil

Prêmios

- 2000** Melhor trabalho da área de Ciências Biológicas e Veterinárias,
Universidade Federal de Minas Gerais

Publicações

- Fogaça, A.C., Zaini, P.A., Wulff, N.A., da Silva, P.I.P., Fázio, M.A., Miranda, A., Daffre, S. & da Silva, A.M. 2010. Effects of the antimicrobial peptide gomesin on the global gene expression profile, virulence, and biofilm formation of *Xylella fastidiosa*. em revisão, FEMS Microbiology Letters
- da Silva, P.I.P., Fogaça, A.C., Zaini, P.A., Teixeira, D.C. & da Silva, A.M. 2010. Global gene expression profile of *Xylella fastidiosa* challenged with sublethal concentration of streptomycin. Em preparação.

Comunicações em Congressos

- Pessoa, P.I. & da Silva, A. M. 2008. Effect of sublethal concentration of streptomycin on *Xylella fastidiosa* gene expression. International Conference on Biotic Plant Interactions, University of Queensland, Austrália
- Pessoa, P.I., Fogaça, A. C. & da Silva, A. M. 2007. Effect of sublethal concentrations of streptomycin in *Xylella fastidiosa*. XXXVI Reunião Annual da Sociedade Brasileira de Bioquímica e Biologia Molecular e IUBMB Conference.

- Pessoa, P.I., Fogaça, A. C. & da Silva, A. M. 2006. Global gene expression profile of *Xylella fastidiosa* in response to sublethal concentrations of streptomycin. 52ª Reunião Anual da Sociedade Brasileira de Genética.
- Pessoa, P.I., Ferreira, E. A. G., Goulart, M. F., Souza. M. B. G., Ribeiro, B. A., Ferreira, R., Álvares, E. & Gomes, F. T. 2000 Comparação da fauna de artrópodes entre duas cavernas com diferente grau de ação antrópica. IX Semana de Iniciação Científica da Universidade Federal de Minas Gerais.
- Pessoa, P.I., Ribeiro, B. A., Goulart, M. F., Souza. M. B. G., Marini, M., Wilson, G. 2000. Comparação de comunidades de invertebrados em três ambientes distintos da estação ecológica do campus da UFMG., XXIII Congresso Brasileiro de Zoologia.
- Pessoa, P.I. 1999. Geração de mutantes do Fator de Transcrição TEAD-1 que perderam CIS-repressão., 13ª Jornada de Biologia e 1º Encontro de Biólogos do CRBio 4ª Região, Pontifícia Universidade Católica de Minas Gerais
- Pessoa, P.I. 1999. Geração de uma biblioteca de mutantes do fator de transcrição TEAD-1 para estudo de domínios funcionais. 45º Congresso Nacional de Genética.

Outras Informações

- Co-orientação de Maurício Barlera. Iniciação científica PIBIC-CNPq, Instituto de Química - USP Projeto: Construção de uma biblioteca genômica de *Xylella fastidiosa*. 2007.
- Ministrante no III Curso de Inverno (2008) do Departamento de Bioquímica do IQ-USP 2008.